

Bildverarbeitung für die Medizin 2010

Veranstalter

IMI	Institut für Medizinische Informatik Rheinisch-Westfälische Technische Hochschule (RWTH) Aachen
JARA-BRAIN	Jülich Aachen Research Alliance, Translational Brain Medicine
BVMI	Berufsverband Medizinischer Informatiker Fachgruppe Medizinische Informatik der
DGBMT	Deutschen Gesellschaft für Biomedizinische Technik im VDE Gemeinsame AG Medizinische Bild- und Signalverarbeitung der
GI	Gesellschaft für Informatik und der
GMDS	Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie
IEEE	Joint Chapter Engineering in Medicine and Biology, German Section

Unterstützende Fachgesellschaften

DAGM	Deutsche Arbeitsgemeinschaft für Mustererkennung
DGaO	Deutsche Gesellschaft für angewandte Optik

Tagungsvorsitz

Prof. Dr. Thomas M. Deserno, geb. Lehmann
Institut für Medizinische Informatik, RWTH Aachen

Tagungssekretariat

Michaela Huth
Institut für Medizinische Informatik, RWTH Aachen
Postanschrift: 52057 Aachen; Lieferanschrift: Pauwelsstr. 30, 52074 Aachen
Telefon: 0241-8088790; Telefax: 0241-3388790
Email: bvm2010@mi.rwth-aachen.de; Web: <http://bvm-workshop.org>

Lokale Organisation

Prof. Til Aach, RWTH Aachen, Fakultät 6
Prof. Katrin Amunts, RWTH Aachen / FZ Jülich GmbH
Prof. Thomas Deserno, RWTH Aachen, Fakultät 10
Prof. Walter Hillen, FH Aachen, Fachbereich 9
Prof. Torsten Kuhlen, RWTH Aachen, Fakultät 1
Prof. Ingrid Scholl, FH Aachen, Fachbereich 5

Verteilte BVM-Organisation

Prof. Dr. Thomas M. Deserno, Jens Hoffmann
Rheinisch-Westfälische Technische Hochschule Aachen (Tagungsband)

Prof. Dr. Heinz Handels, Dipl.-Ing. Martin Riemer
Universitätsklinikum Hamburg-Eppendorf (Begutachtung)

Prof. Dr. Hans-Peter Meinzer, Dipl. Inform. Alexander Seitel
Deutsches Krebsforschungszentrum Heidelberg (Anmeldung)

Prof. Dr. Thomas Tolxdorff, Dagmar Stiller
Charité – Universitätsmedizin Berlin (Internetpräsenz)

Programmkomitee

Prof. Dr. Til Aach, Rheinisch-Westfälische Technische Hochschule Aachen
Prof. Dr. Katrin Amunts, Rheinisch-Westfälische Technische Hochschule Aachen
Prof. Dr. Dirk Bartz, Universität Leipzig
Prof. Dr. Dr. Johannes Bernarding, Universität Magdeburg
Dr. Katja Bühler, VRVis Zentrum für Virtual Reality und Visualisierung Wien
Prof. Dr. Thorsten M. Buzug, Universität zu Lübeck
Prof. Dr. Thomas M. Deserno, Rheinisch-Westfälische Technische Hochschule
Aachen

Prof. Dr. Hartmut Dickhaus, Universität Heidelberg
Dr. Jan Ehrhardt, Universitätsklinikum Hamburg-Eppendorf
Prof. Dr. Dr. Karl-Hans Englmeier, Helmholtz-Zentrum München
Prof. Dr. Rudolf Fahlbusch, International Neuroscience Institute Hannover
Prof. Dr. Bernd Fischer, Universität zu Lübeck
Prof. Dr. Heinz Handels, Universitätsklinikum Hamburg-Eppendorf
Priv.-Doz. Dr. Peter Hastreiter, Universität Erlangen-Nürnberg
Prof. Dr. Hans Herzog, Forschungszentrum Jülich
Prof. Dr. Walter Hillen, Fachhochschule Aachen
Prof. Dr. Joachim Hornegger, Universität Erlangen-Nürnberg
Prof. Dr. Alexander Horsch, Technische Universität München
& Universität Tromsø, Norwegen

Prof. Dr. Justus Ilgner, Rheinisch-Westfälische Technische Hochschule Aachen
Prof. Dr. Dr. Hans-Gerd Lipinski, Fachhochschule Dortmund
Prof. Dr. Gabriele A. Krombach, Rheinisch-Westfälische Technische Hochschule
Aachen

Prof. Dr. Torsten Kuhlen, Rheinisch-Westfälische Technische Hochschule
Aachen

Prof. Dr. Hans-Peter Meinzer, Deutsches Krebsforschungszentrum Heidelberg
Jun.-Prof. Dr. Dorit Merhof, Universität Konstanz
Prof. Dr. Heinrich Müller, Universität Dortmund
Prof. Dr. Henning Müller, Université Sierre, Schweiz
Prof. Dr. Nassir Navab, Technische Universität München

Prof. Dr. Christiane Neuschaefer-Rube, Rheinisch-Westfälische Technische
Hochschule Aachen

Prof. em. Dr. Heinrich Niemann, Universität Erlangen-Nürnberg

Dr. Christoph Palm, Forschungszentrum Jülich

Prof. Dr. Dr. Siegfried J. Pöppel, Universität zu Lübeck

Prof. Dr. Regina Pohle-Fröhlich, Hochschule Niederrhein

Prof. Dr. Bernhard Preim, Universität Magdeburg

Prof. Dr. Karl Rohr, Universität & Dt. Krebsforschungszentrum Heidelberg

Prof. Dr. Peter Rossmanith, Rheinisch-Westfälische Technische Hochschule
Aachen

Prof. Dr. Dietmar Saupe, Universität Konstanz

Dr. Hanno Schar, Forschungszentrum Jülich

Prof. Dr. Ingrid Scholl, Fachhochschule Aachen

Prof. Dr. Nadim J. Shah, Forschungszentrum Jülich

Prof. Dr. Thomas Tolxdorff, Charité – Universitätsmedizin Berlin

Dr. Gudrun Wagenknecht, Forschungszentrum Jülich

Prof. Dr. Rüdiger Westermann, Technische Universität München

Prof. Dr. Herbert Witte, Universität Jena

Dr. Thomas Wittenberg, Fraunhofer IIS Erlangen

Prof. Dr. Ivo Wolf, Hochschule Mannheim

Sponsoren des Workshops BVM 2010

Die BVM wäre ohne die finanzielle Unterstützung der Industrie in ihrer so erfolgreichen Konzeption nicht durchführbar. Deshalb freuen wir uns sehr über langjährige kontinuierliche Unterstützung mancher Firmen sowie auch über das neue Engagement anderer. Dies gilt insbesondere für unsere Platin-Sponsoren.

Platin-Sponsoren

Agfa HealthCare

Konrad-Zuse-Platz 1-3, D-53227 Bonn
<http://www.agfa.com/healthcare>

Siemens AG, Healthcare Sector

Henkestr. 127, D-91052 Erlangen
<http://medical.siemens.com>

Sponsoren

3Dimes GmbH

Daimlerstr. 32, D-60314 Frankfurt/Main, <http://www.3dimes.de>

Barco GmbH

Greschbachstr. 2-4, D-76229 Karlsruhe, <http://www.barco.com>

Cerner Deutschland GmbH

Cunoweg 1, D-65510 Idstein, <http://www.cerner.com>

EST Engineering Systems Technologies GmbH & Co. KG

Postfach 16 50, D-67605 Kaiserslautern, <http://www.est-kl.com>

MeVis Medical Solutions AG

Universitätsallee 29, D-28359 Bremen, <http://www.mevis.de>

NDI Europe GmbH

Fritz-Reichle-Ring 2, D-78315 Radolfzell, <http://www.ndigital.com/>

Philips Technologie GmbH Forschungslaboratorien

Weisshausstr. 2, D-52066 Aachen, <http://www.philips.de>

VISCON GmbH

Inneboltstr. 105, D-47506 Neukirchen-Vluyn, <http://www.viscon.de>

H.F. Wiebe GmbH & Co. KG

Im Finigen 8, D-28832 Achim, <http://www.wiebe.de/>

Stiftung von Preisgeldern

CHILI GmbH Digital Radiology

Burgstr. 61, D-69121 Heidelberg, <http://www.chili-radiology.com>

Springer Science & Business Media Deutschland GmbH

Heidelberger Platz 3, D-14197 Berlin, <http://www.springer.com>

Preisträger des BVM-Workshops 2009 in Heidelberg

Auf der BVM 2009 wurde der mit 1.000 € dotierte BVM-Award an eine herausragende Diplom-, Bachelor-, Master- oder Doktorarbeit aus dem Bereich der Medizinischen Bildverarbeitung vergeben. Die mit einem Preisgeld von je 250 € dotierten BVM-Preise zeichnen besonders hervorragende Arbeiten aus, die auf dem Workshop präsentiert wurden.

BVM-Award 2009 für seine herausragende Dissertation

Matthias Baumhauer (Abteilung Medizinische und Biologische Informatik, DKFZ Heidelberg) Real-time Compensation of Organ Motion for Augmented Reality in Laparoscopic Surgery

BVM-Preis 2009 für die beste wissenschaftliche Arbeit

1. *Jan Paulus* (Lehrstuhl für Mustererkennung, Universität Erlangen-Nürnberg) Automatische Qualitätsmessung von Retina-Fundusbildern
2. *William Godinez* (Forschungsgruppe Bioquant, Universität Heidelberg) Evaluation of Approaches for Tracking Virus Particles in Fluorescence Microscopy Images

BVM-Preis 2009 für den besten Vortrag

1. *Stefan Heldmann* (Institut für Mathematik, Universität zu Lübeck) A Scale-Space Approach for Image Registration of Vessel Structures
2. *Janine Olesch* (Institut für Mathematik, Universität zu Lübeck) Landmark constrained non-parametric image registration with isotropic tolerances

BVM-Preis 2009 für die beste Posterpräsentation

Ivo Rössling (Institut für Simulation und Grafik, Otto-von-Guericke Universität Magdeburg) Interaktive Visualisierung von Abständen und Ausdehnungen Anatomischer Strukturen für die Interventionsplanung

Vorwort

Der Workshop Bildverarbeitung für die Medizin (BVM) hat eine lange Tradition. Er ist im Jahre 1998 aus dem Aachener Workshop, der erstmalig 1996 stattfand, und dem davor schon etablierten Workshop, der jeweils in Freiburg stattfand, hervorgegangen. 1998 erschienen die Proceedings zum Workshop erstmalig in der Reihe Informatik Aktuell des Springer-Verlages. Damals wurde die verteilte Organisation etabliert und auch die jährliche Ausrichtung an wechselnden Standorten festgelegt. Seither ist die BVM in Aachen, Berlin, Erlangen, Hamburg, Heidelberg, Leipzig, Lübeck und München gewesen und qualitativ wie quantitativ ständig gewachsen.

Mittlerweile treffen sich ca. 250 Fachleute aus Wissenschaft, Industrie und Gesundheitsversorgung, um aktuelle Forschungsergebnisse auszutauschen. Ausdrücklich sind auch junge Nachwuchswissenschaftler willkommen, die über ihre Bachelor-, Master-, Diplom- oder Doktorarbeit im Bereich der Medizinischen Bildverarbeitung berichten möchten. Auch die europäischen Nachbarn sind mittlerweile Bestandteil der BVM geworden: Deutsch und Englisch sind gleichberechtigte Tagungssprachen.

Vor diesem Hintergrund freut es uns besonders, dass die BVM nach langer Zeit wieder einmal in Aachen stattfindet. Die Idee eines gemeinsamen Workshops über die Grenzen einzelner Fachgesellschaften und Interessengruppen hinaus, die vor über 10 Jahren hier in Aachen geboren wurde, wurde dieses Jahr besonders ausgelebt: auch die lokale Organisation vereinte die verschiedenen Fakultäten und Fachbereiche der RWTH und der FH Aachen sowie des Forschungszentrums Jülich.

Mit diesem dezentralen Team ist es gelungen, ein äußerst interessantes und vielfältiges Programm zusammenzustellen. Aus 115 Einreichungen wurden 92 Beiträge ausgewählt. Auf der BVM werden 40 Vorträge, vier wissenschaftliche Software-Demonstrationen und 48 Poster präsentiert. Zwei eingeladene Vorträge und drei Tutorien runden das Wissenschaftliche Programm ab:

- V1 Prof. Bernd Jähne (Heidelberg):
Robuste Merkmale und Segmentierungsverfahren: neue Ansätze
- V2 Prof. Axel W. E. Wiesmüller (Rochester, New York, USA):
Pendeln zwischen den Welten: Bildverarbeitung oder Medizin? in Deutschland versus USA
- T1 Prof. Ivo Wolf (Mannheim) & Dr. Marco Nolden (Heidelberg):
Toolkits für medizinische Applikationen
- T2 Prof. Berthold Wein (Aachen):
Bildverarbeitung in der Radiologischen Praxis
- T3 Dipl.-Inform. Michael Onken (Oldenburg):
Digital Imaging and Communications in Medicine (DICOM) und Integrating the Healthcare Enterprises (IHE)

Am Montagabend findet ein gemeinsamer Gesellschaftsabend im historischen Tonnengewölbe des gotischen Rathauses zu Aachen statt, das auf den Grundmauern der Palastaula der karolingischen Kaiserpfalz im 14. Jahrhundert von

der Aachener Bürgerschaft erbaut wurde und seit 1978 mit zum Weltkulturerbe der UNESCO zählt. Der Granusturm des Rathauses zeugt noch heute von der ursprünglichen Bebauung aus der Zeit Karls des Großen.

Vor diesem Hintergrund wünschen die Herausgeber allen Teilnehmerinnen und Teilnehmern eine interessante, lebhafte und erfolgreiche Tagung.

Außerdem möchten wir allen herzlich danken, die zum Gelingen des BVM-Workshops 2010 beigetragen haben: den Mitarbeiterinnen und Mitarbeitern des Instituts für Medizinische Informatik der RWTH Aachen für die Organisation vor Ort, insbesondere Frau Michaela Huth, die sich mit ausserordentlichem Engagement um die BVM gekümmert und das Tagungssekretariat vorbildlich geleitet hat, den Herren Detlef Jung und Martin Bodych für die technische Unterstützung, sowie Frau Marion Fuchs und den Herren Sebastian Thoennes und Max Jung.

Wir danken Frau Dagmar Stiller für die Pflege der Internetpräsentation, Herrn Alexander Seitel für die Pflege des BVM-Emailverteilers und der webgestützten Workshopanmeldung, Herrn Martin Riemer für die web-basierte Durchführung des Reviewing-Prozesses und Herrn Jens Hoffmann für die Überarbeitung der L^AT_EX-Styles sowie für die Entwicklung der web-basierten Beitragseinreichung. Frau Dorothea Glaunsinger vom Springer-Verlag danken wir für die langjährige gute Kooperation. Schließlich danken wir noch einmal allen Sponsoren, ohne deren finanzielle Unterstützung der BVM-Workshop in seiner etablierten Form nicht durchführbar wäre.

Aachen im Januar 2010

Thomas Deserno
Heinz Handels
Hans-Peter Meinzer
Thomas Tolxdorff

Inhaltsverzeichnis

Die fortlaufende Nummer am linken Seitenrand entspricht den Beitragsnummern, wie sie im endgültigen Programm des Workshops zu finden sind. Dabei steht V für Vortrag, P für Poster und S für Software demonstration.

Bildgebung Modalitäten

V1	<i>Knopp T, Biederer S, Sattel TF, Weizenecker J, Gleich B, Borgert J, Buzug TM: Rekonstruktion von Magnetic Particle Imaging Daten mittels einer modellierten Systemfunktion</i>	1
V2	<i>Biederer S, Sattel TF, Knopp T, Buzug TM: Variable Trajektoriendichte für Magnetic Particle Imaging</i>	6
V3	<i>Seitel A, dos Santos TR, Mersmann S, Penne J, Tetzlaff R, Meinzer H-P, Maier-Hein L: Time-of-Flight Kameras für die intraoperative Oberflächenerfassung</i>	11

Navigation

V4	<i>Penne J, Schaller C, Engelbrecht R, Maier-Hein L, Schmauss B, Meinzer H-P, Hornegger J: Laparoscopic Quantitative 3D Endoscopy for Image Guided Surgery</i>	16
V5	<i>Gergel I, dos Santos TR, Tetzlaff R, Maier-Hein L, Meinzer H-P, Wegner I: Partikelfilterung für die Kompensation von Atembewegung während der navigierten Bronchoskopie</i>	21
V6	<i>Martin-Gonzalez A, Heining SM, Navab N: Sight-based Magnification System for Surgical Applications</i>	26

Registrierung

V7	<i>Werner R, Wolf -J-C, Ehrhardt J, Schmidt-Richberg A, Handels H: Automatische Landmarkendetektion und -übertragung zur Evaluation der Registrierung von thorakalen CT-Daten</i>	31
----	---	----

V8	<i>dos Santos TR, Gergel I, Mersmann S, Meinzer H-P, Maier-Hein L:</i> Graphbasierte Registrierung von Tubulären Strukturen	36
V9	<i>Behrens A, Bommers M, Stehle T, Gross S, Leonhardt S, Aach T:</i> Mosaickingalgorithmus zur schnellen Panoramabilderstellung in der Fluoreszenzendoskopie	41

Visualisierung

V10	<i>Mühler K, Preim B:</i> Günstige Kamerapfade für medizinische Animationen	46
V11	<i>Behrens A, Guski M, Stehle T, Gross S, Aach T:</i> Intensitätsbasiertes Multiskalen-Blending zur Erstellung von Panoramabildern in der Fluoreszenzendoskopie	51
V12	<i>Würfel W, Hussong A, Herzog A, Erfurt P, Majdani O, Rau TS:</i> Verfahren zur hochgenauen 3D-Rekonstruktion aus histologischen Schliffbildern	56

Bildanalyse

V13	<i>Valentinitsch A, Patsch J, Mueller D, Kainberger F, Langs G:</i> Texture Analysis in Quantitative Osteoporosis Assessment	61
V14	<i>Säring D, Müllerleile K, Groth M, Lund G, Handels H:</i> Lokale Analyse von Infarkttrandzonen in 3D-DE-MRT Bildsequenzen	66
V15	<i>Wagner F, Elter M:</i> Merkmale zur Beschreibung der Intensitätsvariation für die Klassifikation von Herdbefunden in Mammogrammen	71

Mikroskopie

V16	<i>Weichert F, Gaspar M, Zybin A, Gurevich EL, Görtz A, Timm C, Müller H, Marwedel P:</i> Plasmonen-unterstützte Mikroskopie zur Detektion von Viren	76
-----	---	----

V17	<i>Harder N, Mora-Bermúdez F, Godinez WJ, Wünsche A, Ellenberg J, Eils R, Rohr K: Automatic Analysis of Live Cell Image Sequences to determine Temporal Mitotic Phenotypes</i>	81
V18	<i>Wörz S, Sander P, Pfannmöller M, Rieker RJ, Joos S, Mechttersheimer G, Boukamp P, Lichter P, Rohr K: Model-Based Segmentation and Colocalization Quantification in 3D Microscopy Images</i>	86

Software demos

S1	<i>Schönmeyer R, Athelougou M, Sittek H, Ellenberg P, Feehan O, Schmidt G, Binnig G: Prototyp eines Mammographie-CAD-Systems auf Basis der Cognition Network Technology</i>	92
S2	<i>Kaster FO, Kassemeyer S, Merkel B, Nix O, Hamprecht FA: An Object-oriented Library for Systematic Training and Comparison of Classifiers for Computer-assisted Tumor Diagnosis from MRSI Measurements</i>	97
S3	<i>Wieczorek M, Aichert A, Kutter O, Bichlmeier C, Landes J, Heining SM, Euler E, Navab N: GPU-accelerated Rendering for Medical Augmented Reality in Minimally-invasive Procedures</i>	102
S4	<i>Emmersberger M, Demirci S, Ghotbi R, Navab N: SASOMI</i>	107

Postersession 1

P1	<i>John C, Schwanecke U: Ein System zur berührungslosen, volumetrischen Vermessung von Gesichtsschwellungen</i>	112
P2	<i>Selby BP, Sakas G, Walter S, Groch W-D, Stilla U: A Radiometry Tolerant Method for Direct 3D/2D Registration of Computed Tomography Data to X-ray Images</i>	117
P3	<i>Kratz B, Oehler M, Buzug TM: Vorwissensbasierte NFFT zur CT-Metallartefaktreduktion</i>	122
P4	<i>Pritzkau A, Bartz D: Parahistogramme innerhalb eines dreidimensionalen Interaktionsraumes</i>	127
P5	<i>Chen L, Bruijns J, ter Romeny BMH: Acceleration of the Fully Automatic Branch Labeling of Voxel Vessel Structures</i>	132

P6	<i>Witte M, Wex C, Riefenstahl N, Michaelis B, Jacob S, Lippert H:</i> Photogrammetrische 3D-Vermessung von Organen	137
P7	<i>Born S, Wellein DI, Zöllner A, Bartz D:</i> Segmentation-Enhanced Registration of Angiography Data	142
P8	<i>Merhof D, Greiner G, Buchfelder M, Nimsky C:</i> Fiber Selection from Diffusion Tensor Data based on Boolean Operators	147
P9	<i>Ulrich C, Schaller C, Penne J, Hornegger J:</i> Evaluation of a Time-of-Flight-based Respiratory Motion Management System	152
P10	<i>Simbt S, Dennerlein F, Boese J:</i> Markerbasiertes Online Kalibrierverfahren für die CT-Rekonstruktion	157
P11	<i>Hentschke CM, Tönnies KD:</i> Automatic 2D/3D-Registration of Cerebral DSA Data Sets	162
P12	<i>Kellermann K, Baer A, Preim B:</i> Adaptive Fokus-Kontext-Kategorisierung für Visualisierungen zur Operationsplanung	167
P13	<i>Schäfer S, Tönnies KD:</i> Detection of Motion Distorted Areas in Perfusion MRI of the Breast	172
P14	<i>Placht S, Schaller C, Balda M, Adelt A, Ulrich C, Hornegger J:</i> Improvement and Evaluation of a Time-of-Flight-based Patient Positioning System	177
P15	<i>Walczak L, Weichert F, Schröder A, Landes C, Müller H, Wagner M:</i> Einfluss von Formvariationen auf Finite Elemente Simulationen bei muskulären Strukturen	182
P16	<i>Winter S, Ritschel K, Broll M, Dekomien C:</i> Kalibrierung eines 3D-Ultraschallsystems mit evolutionärer Optimierung	187
P17	<i>Balda M, Heismann BJ, Hornegger J:</i> Non-Stationary CT Image Noise Spectrum Analysis	191
P18	<i>Wörmann J, Braun A, Mempel M, Englmeier KH, Hamm P:</i> Automatisierte quantitative Analyse der Zellzusammensetzung von bronchoalveolaren Spülungen	196
P19	<i>Weirich C, Scheins J, Gaens M, Tellmann L, Kops ER, Kaffanke J, Shah J, Herzog H:</i> Simultaneous PET and MR Imaging with a Newly Developed 3TMR-BrainPET Scanner	201

P20	<i>Gross S, Schink M, Stehle T, Behrens A, Tischendorf J, Trautwein C, Aach T: Echtzeitfähige Extraktion scharfer Standbilder in der Video-Koloskopie</i>	206
P21	<i>Imhäuser C, Gulbins (Grassmé) H, Gulbins E, Lipinski H-G: 3D-Visualisierung und Kolokalisation von Proteinen und ceramidreichen Domänen</i>	211
P22	<i>Engel M, Seitel A, Fangerau M, Redeleff BA, Sommer CM, Essert-Villard C, Baegert C, Meinzer H-P, Maier-Hein L: Schnelle Zugangsplanung für die perkutane Punktion der Leber</i>	216
P23	<i>Paulus J, Bock R, Daum V, Hornegger J: Non-Rigid Registration to Capture Optic Nerve Head Variability</i>	221
P24	<i>Wald D, Schwarz T, Fangerau M, Dinkel J, Delorme S, Kaaks R, Meinzer H-P: Effiziente Methode zur Generierung von Ganzkörperdaten für die Fettgewebsanalyse</i>	226

Algorithmen

V19	<i>Maier-Hein L, dos Santos TR, Franz AM, Meinzer H-P: Iterative Closest Point Algorithm in the Presence of Anisotropic Noise</i>	231
V20	<i>Wang X, Wolf I, Hartmann P, Heimann T, Meinzer H-P, Wegner I: Ein gradientenflussbasiertes Ähnlichkeitsmaß für das Tracking von Gefäßen</i>	236
V21	<i>Wieczorek H: Reconstruction Image Quality Theory</i>	241

Toolkits

V22	<i>Fritzsche K, Meinzer H-P: MITK-DI</i>	246
V23	<i>Fried E, Geng Y, Ullrich S, Kneer D, Grottke O, Rossaint R, Deserno TM, Kuhlen T: MEDOX</i>	251
V24	<i>Hamo O, Nelles G, Wagenknecht G: A Design Toolbox to Generate Complex Phantoms for the Evaluation of Medical Image Processing Algorithms</i>	256

Segmentierung

- V25 *Budai A, Michelson G, Hornegger J*: Multiscale Blood Vessel Segmentation in Retinal Fundus Images 261
- V26 *Greß O, Möller B, Stöhr N, Hüttelmaier S, Posch S*: Scale-adaptive Wavelet-based Particle Detection in Microscopy Images 266
- V27 *Wellein DI, Pfeifle M, Althuizes M, Voitel L, Bartz D*: A Cortex Segmentation Pipeline for Neurosurgical Intervention Planning 271
- V28 *Goößen A, Hermann E, Pralow TGT, Grigat R-R*: Model-Based Lower Limb Segmentation using Weighted Multiple Candidates 276

Modellierung

- V29 *Ruppertshofen H, Lorenz C, Beyerlein P, Salah Z, Rose G, Schramm H*: Fully Automatic Model Creation for Object Localization utilizing the Generalized Hough Transform 281
- V30 *Gollmer ST, Buzug TM*: Statistische 3D Formmodellierung mittels quasi-verzerrungsfreier sphärischer Parametrisierung 286
- V31 *Kirschner M, Wesarg S*: 3D Statistical Shape Model Building using Consistent Parameterization 291
- V32 *Georgi J-C, Bippus R, Wang WW, Lee NY, Narayanan M, Schöder H, Guillem J, Humm JL*: Pharmakokinetische Modellierung von FMISO-PET/CT Bildgebung bei Plattenepithelkarzinomen im Kopf-Halsbereich 296

Postersession 2

- P25 *Eiben B, Palm C, Pietrzyk U, Amunts CDK*: Perspective Error Correction using Registration for Blockface Volume Reconstruction of Serial Histological Sections of the Human Brain 301
- P26 *Becker S, Mang A, Buzug TM*: Approximation des Tumormasseeffekts mittels direkt-manipulierender Free-Form Deformation 306
- P27 *Schipptritt D, Wiemann M, Lipinski H-G*: Bildgestützte Analyse der in vitro-Sedimentation agglomerierter Nanopartikel 311

P28	<i>Fritsche A, Fischer B, Deserno TM</i> : Orts-relationale SIFT-Hierarchien zur Ähnlichkeitsbestimmung mit Graph-Matching	315
P29	<i>Dey T, Wieczorek H, Backus B, Romijn L, Bippus R, Verzijlbergen J, Aach T</i> : Thallium-Stress, Technetium-Rest Protokoll für Cardiac SPECT	320
P30	<i>Jung F, Wesarg S</i> : 3D Registration based on Normalized Mutual Information	325
P31	<i>Forkert ND, Säring D, Eisenbeis A, Leyppoldt F, Fiehler J, Handels H</i> : Experimental Assessment of Infarct Lesion Growth in Mice using Time-Resolved T2* MR Image Sequences	330
P32	<i>Papenberg N, Schumacher H, Heldmann S, Böhler T, van Straaten D, Wirtz S</i> : Multimodale Registrierung von Knochen-Szintigraphien und Röntgenbildern	335
P33	<i>Welter P, Gülpers R, Deserno TM, Eichelberg JRM, Onken M, Grouls C, Günther RW</i> : Entwurf eines DICOM Structured Report am Beispiel Content-Based Image Retrieval	340
P34	<i>Schweizer B, Goedicke A</i> : Validation of GEANT4 for Accurate Modeling of ^{111}In SPECT Acquisition	345
P35	<i>Engelhardt S, Ameling S, Wirth S, Paulus D</i> : Features for Classification of Polyps in Colonoscopy	350
P36	<i>Prochiner AS, Overhoff HM</i> : Determination of a Vessel Tree Topology by Different Skeletonizing Algorithms	355
P37	<i>Overhoff HM, Bußmann S</i> : Online Detection of Straight Lines in 3-D Ultrasound Image Volumes for Image-Guided Needle Navigation	360
P38	<i>Schwemmer C, Prümmer M, Daum V, Hornegger J</i> : High-Density Object Removal from Projection Images using Low-Frequency-Based Object Masking	365
P39	<i>Bippus R, Goedicke A, Botterweck H</i> : Monte-Carlo-Based Scatter Correction for Quantitative SPECT Reconstruction	370
P40	<i>Stoll A, Fränzle A, Bendl R</i> : Data Reduction for Supervised Learning in Medical Image Analysis	375
P41	<i>Hofmann HG, Keck B, Hornegger J</i> : Accelerated C-Arm Reconstruction by Out-of-Projection Prediction	380
P42	<i>Radmer J, Krüger J</i> : Ganganalyse für die klinische Anwendung auf Basis von Tiefendaten	385

P43	<i>Zoehrer F, Drexel J, Hahn H</i> : Speckle Reduction for Automated Breast Ultrasound	390
P44	<i>Rössling I, Hahn P, Dornheim L</i> : Schätzung der Midsagittalebene zur Bestimmung der Seitenlage maligner Strukturen des Halses	395
P45	<i>Scholl I, Schubert N, Ziener P, Pietrzyk U</i> : GPU-basiertes Volumenrendering von multimodalen medizinischen Bilddaten in Echtzeit	400
P46	<i>Siewert R, Schnapauff D, Denecke T, Tolxdorff T, Krefting D</i> : Automatic Liver Segmentation in Contrast-enhanced MRI	405
P47	<i>Reichl T, Kutter O, Schultis B, Menzel M, Hautmann H, Navab N</i> : Video-basiertes Tracking eines Bronchoskops	410
P48	<i>Pommerencke T, Westphal K, Ernst C, Dickhaus H, Grabe N</i> : Image-based Quantification of Skin Irritation by Spatial Biomarker Profiling	415

Interaktive Messungen

V33	<i>Proksch D, Dornheim J, Preim B</i> : Interaktionstechniken zur Korrektur medizinischer 3D-Segmentierungen	420
V34	<i>Forkert ND, Schmidt-Richberg A, Säring D, Fiehler J, Illies T, Möller D, Handels H</i> : Graphen- und Level-Set-basierte Nachverarbeitung von 3D-Gefäßsegmentierungen	425
V35	<i>Teßmann M, Vega-Higuera F, Bischoff B, Hausleiter J, Greiner G</i> : Robust Automatic Calcium Scoring for CT Coronary Angiography .	430
V36	<i>Stehle T, Wulff J, Behrens A, Gross S, Aach T</i> : Modellbasierte Echtzeit-Bewegungsschätzung in der Fluoreszenzendoskopie	435

Medizinische Anwendungen

V37	<i>Rietdorf U, Riesenkampff E, Schwarz T, Kuehne T, Meinzer H-P, Wolf I</i> : Planung und Simulation von Patchimplantaten zur intrakardialen Korrektur angeborener Herzfehler	440
V38	<i>Wucherer P, Bichlmeier C, Eder M, Kovacs L, Navab N</i> : Multimodal Medical Consultation for Improved Patient Education .	445

V39	<i>Ilgner J, Park J-HJ, Westhofen M</i> : Praktische Aspekte zur hochauflösenden Stereovideo-Dokumentation intraoperativer Befunde in der Hals-, Nasen- und Ohrenheilkunde	450
V40	<i>Barthold C, Papst A, Küblbeck TWC, Lautenbacher S, Schmid U, Friedl S</i> : Tracking von Gesichtsmimik mit Hilfe von Gitterstrukturen zur Klassifikation von schmerzrelevanten Action Units	455

Rekonstruktion von Magnetic Particle Imaging Daten mittels einer modellierten Systemfunktion

Tobias Knopp¹, Sven Biederer¹, Timo F. Sattel¹, Jürgen Weizenecker²,
Bernhard Gleich², Jörn Borgert², Thorsten M. Buzug¹

¹Institut für Medizintechnik, Universität zu Lübeck, Lübeck

²Philips Technologie GmbH Forschungslaboratorien, Hamburg

knopp@imt.uni-luebeck.de

Kurzfassung. Magnetic Particle Imaging ist ein neues quantitatives Bildgebungsverfahren, das es erlaubt, die örtliche Verteilung super-paramagnetischer Nanopartikel zu bestimmen. Zur Rekonstruktion wird die Systemfunktion benötigt, die den Zusammenhang zwischen dem Messsignal und der Partikelverteilung beschreibt. Hierzu wird bislang eine zeitintensive Kalibrierungsmessung verwendet, bei der eine kleine Delta-probe durch das Messvolumen bewegt und die Partikelantwort an jedem Ort gemessen wird. In dieser Arbeit wird gezeigt, dass statt dieser Kalibrierungsmessung auch ein Modell der Signalkette genutzt werden kann, um die Systemfunktion zu ermitteln. Die Genauigkeit der modellierten Systemfunktion wird anhand von rekonstruierten 2D-MPI-Bildern evaluiert.

1 Einleitung

Das bildgebende Verfahren Magnetic Particle Imaging (MPI) ermöglicht es, die örtliche Verteilung von super-paramagnetische Eisenoxid-Nanopartikeln (sogenannten SPIOs) zu bestimmen. Es wurde von Gleich und Weizenecker entwickelt und zum ersten Mal in [1] veröffentlicht. Seitdem wurde der erste experimentelle Aufbau weiterentwickelt, so dass die Echtzeitfähigkeit von MPI gezeigt werden konnte [2]. In [3] wurden die ersten 3D *in-vivo*-Aufnahmen veröffentlicht, die Details von einem schlagenden Mäuseherzen zeigen.

Zur Bildgebung nutzt MPI die nichtlineare Magnetisierungskurve der Nanopartikel. Werden diese mit einem sinusförmigen Wechselfeld angeregt, so hat die zeitliche Änderung der Magnetisierung näherungsweise die Form einer Rechteckfunktion. Mit einer Empfangsspule kann diese Magnetisierungsänderung gemessen werden. Aufgrund der Nichtlinearität enthält das Spektrum der induzierten Spannung nicht nur die Anregungsfrequenz sondern auch harmonische Frequenzen, die für die Bildgebung genutzt werden können. Zur Ortskodierung verwendet MPI ein statisches Gradientenfeld, dessen Feldstärke nur an einem bestimmten Ort - dem feldfreien Punkt (FFP) - gleich Null ist und ansonsten linear im Raum ansteigt.

Der lineare Zusammenhang zwischen dem Empfangssignal und der Partikelverteilung kann durch die Systemfunktion beschrieben werden, welche die

Partikeldynamik, die Spulengeometrie und die Messparameter enthält [1]. Zur Rekonstruktion der Partikelverteilung muss die Systemfunktion bekannt sein. Hierzu wird bislang eine zeitintensive Kalibrierungsmessung durchgeführt, die eine Deltaprobe nutzt, welche mit den später in der Anwendung verwendeten Nanopartikeln gefüllt ist. Die Deltaprobe repräsentiert einen Voxel und wird zu jedem Ortspunkt bewegt, an dem die Partikelkonzentration rekonstruiert werden soll. Dabei wird jeweils eine Messung durchgeführt, so dass die Systemfunktion sukzessive ermittelt werden kann.

Obwohl dieses messbasierte Verfahren zu guten Ergebnissen führt [1, 3, 4], hat es einige Nachteile. Zum einen ist die gemessene Systemfunktion rauschbehaftet, wodurch die Bildqualität beeinträchtigt wird. Da das Signal-zu-Rausch-Verhältnis (SNR) von der Größe der Deltaprobe abhängt, ist die Bildauflösung bei dem Verfahren beschränkt. Zum anderen ist das Verfahren sehr zeitaufwendig und benötigt für ein 3D-Gitter der Größe $34 \times 20 \times 28$ ca. 6 Stunden Messzeit [3].

Kürzlich wurde ein alternatives Verfahren vorgeschlagen [5], das die Systemfunktion auf Grundlage eines physikalischen Modells der Signalkette berechnet. Erste 1D-Ergebnisse zeigen, dass eine modellierte Systemfunktion gegenüber einer gemessenen vergleichbare Ergebnisse liefert. In dieser Arbeit wird zum ersten Mal die Genauigkeit der modellbasierten Rekonstruktion für 2D-MPI-Daten untersucht.

2 Material and Methoden

Bei der 2D-MPI-Bildgebung werden in zwei Empfangsspulen Spannungssignale $u_l(t)$, $l = 0, 1$, $t \in [0, T]$ gemessen. Der Zusammenhang zwischen der Partikelkonzentration c an der Position \mathbf{r} und der Frequenzkomponente $\hat{u}_{l,k} = \int_0^T u_l(t) e^{2\pi i k t / T} dt$ kann durch

$$\hat{u}_{l,k} = \int_{\Omega} \hat{s}_{l,k}(\mathbf{r}) c(\mathbf{r}) d\mathbf{r} \quad (1)$$

beschrieben werden. Dabei bezeichnet Ω das Volumen, in dem sich Nanopartikel befinden und $\hat{s}_{l,k}(\mathbf{r})$ die Systemfunktion. Wie in [5] vorgeschlagen, kann die Systemfunktion durch

$$\hat{s}_{l,k}(\mathbf{r}) = -\hat{a}_{l,k} \mu_0 \mathbf{p}_l^R(\mathbf{r}) \cdot \int_0^T \frac{\partial}{\partial t} \widehat{\mathbf{M}}(\mathbf{H}(\mathbf{r}, t)) e^{-2\pi i k t / T} dt \quad (2)$$

modelliert werden. Dabei bezeichnet $\mathbf{H}(\mathbf{r}, t)$ die orts- und zeitabhängige magnetische Feldstärke, $\widehat{\mathbf{M}}(\mathbf{H}(\mathbf{r}, t))$ die Partikelmagnetisierung bei Einheitskonzentration, $\mathbf{p}_l^R(\mathbf{r})$ die Sensitivität der l -ten Empfangsspule, $\hat{a}_{l,k}$ die Übertragungsfunktion des l -ten Empfangspfads, T die Messzeit und μ_0 die Permeabilität im Vakuum.

Zur Berechnung der magnetische Feldstärke und der Empfangsspulensensitivitäten wird das Biot-Savart-Gesetz verwendet. Hierzu ist es notwendig, die

genaue Geometrie aller felderzeugenden Elemente zu kennen. Die Partikelmagnetisierung kann durch die Langevin-Theorie des Paramagnetismus modelliert werden [6]. Einzig für die Übertragungsfunktion wird eine kurze Kalibrierungsmessung verwendet. Dazu wird, wie in [5] beschrieben, die Systemfunktion an einigen wenigen Ortspunkten gemessen und die Übertragungsfunktion mittels eines Kleinste-Quadrate-Ansatzes optimiert. In dieser Arbeit werden hierzu 24 Ortspunkte verwendet, so dass die Kalibrierungsmessung 15 Sekunden benötigt.

Als Messaufbau wird der MPI-Scanner verwendet, der in [3] vorgestellt wurde. Eine Übersicht über die Scanner-Geometrie ist in Abb. 1 gegeben. Alle geometrischen Parameter wurden vermessen, um die magnetische Feldstärke simulieren zu können. Das verwendete Phantom besteht aus 12 Bohrungen mit einem Durchmesser von 0.5 mm und einer Länge von 1 mm und stellt den Buchstaben „P“ dar (Abb. 1). Die Bohrungen sind mit unverdünntem Resovist[®] (Bayer Schering Pharma AG) gefüllt, einem kommerziell erhältlichen SPIO Kontrastmittel, das üblicherweise in der Magnetresonanztomographie genutzt wird.

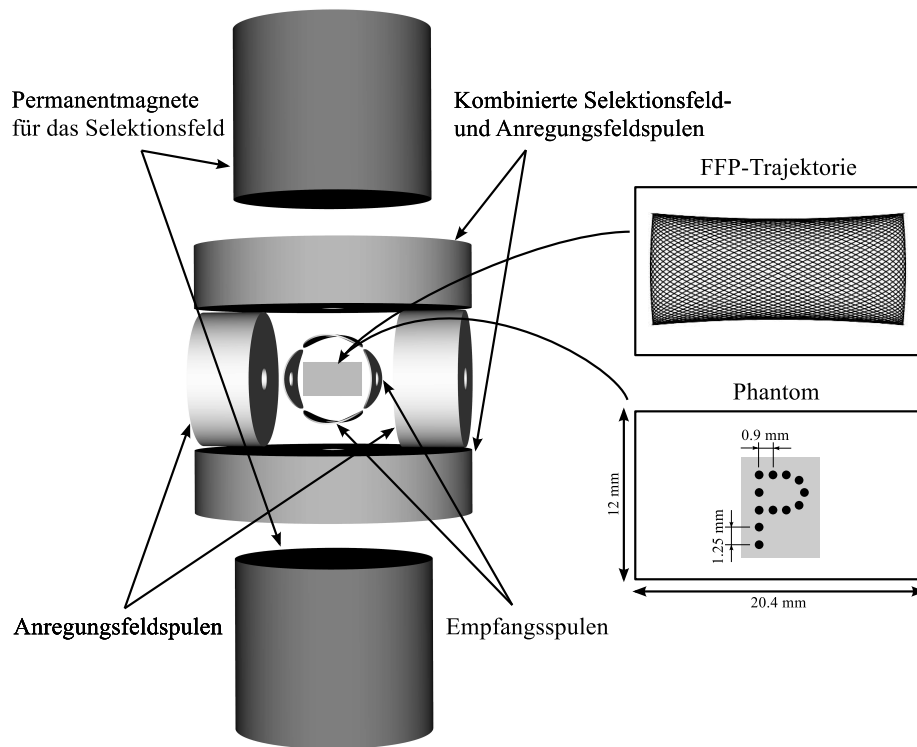


Abb. 1. Aufbau des 2D-MPI-Scanners, der aus zwei Permanentmagneten, zwei Sendespulenpaaren, und zwei Empfangsspulenpaaren besteht. Des Weiteren ist die Lissajous Abtasttrajektorie und das verwendete Phantom gezeigt.

Als Referenz für die modellierte Systemfunktion wird eine gemessene Systemfunktion genutzt. Hierzu wird eine Deltaprobe in der Form eines Würfels der Größe $(0.6 \text{ mm})^3$, gefüllt mit unverdünntem Resovist[®], durch den Messbereich gefahren. Die Systemfunktion wird an 2720 Positionen innerhalb eines Messbereichs von $20,4 \text{ mm} \times 12,0 \text{ mm}$ auf einem Gitter der Größe 68×40 ausgewertet. Diese Kalibrierungsmessung benötigt 45 Minuten. Neben der gemessenen werden zwei modellierte Systemfunktionen genutzt. Die erste verwendet dieselbe Gittergröße wie die gemessene Systemfunktion. Die zweite nutzt ein Gitter mit doppelter Auflösung (136×80). Während die modellierte Systemfunktion auf einem beliebigen Gitter ausgewertet werden kann, ist die Gittergröße bei der gemessenen Systemfunktion beschränkt, da das Messsignal und somit auch die Systemfunktion umso verrauschter ist, je kleiner die Deltaprobe gewählt wird.

Zur Rekonstruktion wird das Gleichungssystem, das durch Diskretisierung von (1) entsteht, mittels eines regularisierten Kleinste-Quadrate-Ansatzes gelöst. Hierzu wird der Kaczmarz-Algorithmus verwendet, der bei MPI eine hohe Konvergenzgeschwindigkeit aufweist [3].

3 Ergebnisse

In Abb. 2 sind die Rekonstruktionsergebnisse einer dynamischen Bildsequenz an vier verschiedenen Zeitpunkten gezeigt. Die Ergebnisse bei Verwendung der modellierten Systemfunktion sind in großen Bildbereichen vergleichbar mit denen bei Verwendung der gemessenen Systemfunktion. Lediglich am Bildrand sind die Ergebnisse, die mit der gemessenen Systemfunktion rekonstruiert wurden,

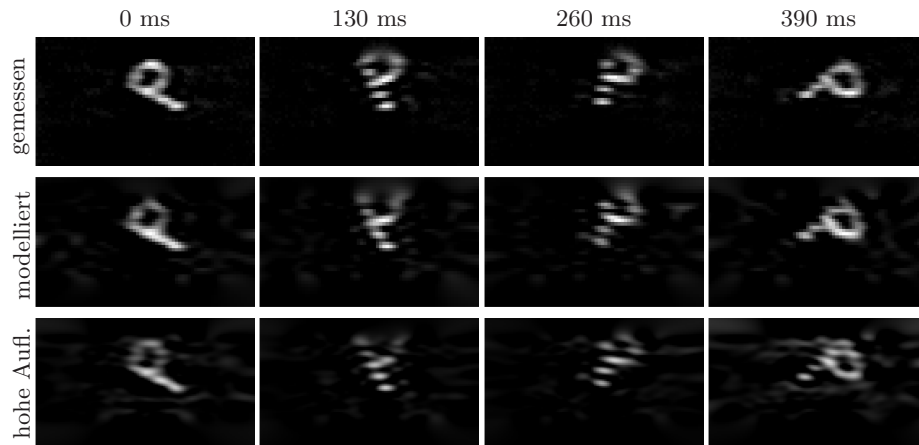


Abb. 2. Ergebnisse der Rekonstruktion mit unterschiedlichen Systemfunktionen an vier ausgewählten Zeitpunkten. Verwendet wurden eine gemessene Systemfunktion (Gittergröße 68×40) und zwei modellbasierte Systemfunktionen (Gittergrößen 68×40 und 136×80).

besser. Die Bilder, die mit der hochaufgelösten Systemfunktion rekonstruiert wurden, zeigen zusätzliche Artefakte. Allerdings können im ersten und letzten Bild der Sequenz mehr Punkte unterschieden werden als bei der Verwendung der gemessenen Systemfunktion. Diese Auflösungsverbesserung macht das Potential der modellbasierten Rekonstruktion deutlich.

4 Diskussion

Die Rekonstruktionsergebnisse zeigen, dass eine modellierte Systemfunktion nahezu vergleichbare Ergebnisse wie eine gemessene Systemfunktion liefert. Die modellierte Systemfunktion wurde inklusive Kalibrierungsmessung für die Übertragungsfunktion in nur 15 Sekunden bestimmt, während die gemessene Systemfunktion 45 Minuten benötigte. Wenn man bedenkt, dass eine 3D-Systemfunktion auf einem 128^3 (256^3)-Gitter einen Monat (sechs Monate) Messzeit benötigt, wird deutlich, wie wichtig eine modellbasierte Rekonstruktion für MPI ist.

Um das auf einer Kalibrierungsmessung basierte Verfahren endgültig abzulösen, muss die Genauigkeit der modellierten Systemfunktion allerdings noch verbessert werden. Vermutlich ist das genutzte Partikelmodell zu ungenau, da es keine Anisotropien der Partikel berücksichtigt [7]. Bei einem direkten Vergleich der modellierten und der gemessenen Systemfunktion haben wir weiter festgestellt, dass die gemessene Systemfunktion geometrische Verzerrungen am Bildrand aufweist, die bei der modellierten Systemfunktion nicht auftreten. Eine Möglichkeit, diese Verzerrungen näher zu untersuchen, ist die Verwendung einer Bildregistrierung zwischen der gemessenen und der modellierten Systemfunktion. Durch Einführen eines geeigneten Verrückungsfelds, könnte die Genauigkeit der modellbasierten Systemfunktion weiter verbessert werden.

Literaturverzeichnis

1. Gleich B, Weizenecker J. Tomographic imaging using the nonlinear response of magnetic particles. *Nature*. 2005;435(7046):1214–7.
2. Gleich B, Weizenecker J, Borgert J. Experimental results on fast 2D-encoded magnetic particle imaging. *Phys Med Biol*. 2008;53(6):N81–N84.
3. Weizenecker J, Gleich B, Rahmer J, et al. Three-dimensional real-time in vivo magnetic particle imaging. *Phys Med Biol*. 2009;54(5):L1–L10.
4. Sattel T, Knopp T, Biederer S, et al. Single-Sided device for magnetic particle imaging. *J Phys D Appl Phys*. 2009;42(1):1–5.
5. Knopp T, Sattel T, Biederer S, et al. Model-based reconstruction for magnetic particle imaging. *IEEE Trans Med Imag*. 2009;published online.
6. Chikazumi S, Charap S. *Physics of Magnetism*. New York: Wiley; 1964.
7. Biederer S, Knopp T, Sattel T, et al. Magnetization response spectroscopy of superparamagnetic nanoparticles for magnetic particle imaging. *J Phys D Appl Phys*. 2009;42(20):7pp.

Variable Trajektoriendichte für Magnetic Particle Imaging

Sven Biederer, Timo F. Sattel, Tobias Knopp, Thorsten M. Buzug

Institut für Medizintechnik, Universität zu Lübeck, Lübeck
biederer@imt.uni-luebeck.de

Kurzfassung. Bei dem bildgebenden Verfahren Magnetic Particle Imaging (MPI) wird die Ortskodierung dadurch erzielt, dass ein sogenannter feldfreier Punkt (FFP) auf einer Trajektorie durch das Betrachtungsfeld gefahren wird. Als eine der besten Möglichkeiten haben sich dafür Lissajous-Figuren erwiesen. Ein Nachteil ist dabei, dass in einem MPI-System die Hardware auf feste Frequenzen abgestimmt werden muss. Somit entsteht eine statische Trajektorie, die sowohl eine feste örtliche Auflösung als auch eine feste Aufnahmezeit besitzt. In diesem Beitrag wird eine Variation der Phase vorgestellt, die es ermöglicht, die örtliche Auflösung und Aufnahmezeit zu verändern. Dies ermöglicht zum Beispiel einen schnellen Orientierungsscan mit niedriger Bildqualität oder einen hochauflösenden Detailscan mit längerer Aufnahmezeit.

1 Einleitung

Magnetic Particle Imaging (MPI) ist ein neues bildgebendes Verfahren, das die örtliche Verteilung von superparamagnetischen Eisenoxid-Nanopartikeln (SPIOs) misst. Bei der quantitativen Darstellung dieser Nanopartikel erzielt MPI eine örtliche Auflösung im Submillimeterbereich. Die Aufnahmezeit ist dabei so kurz, dass Bilder in Echtzeit aufgenommen werden können. Erstmals wurde MPI von Gleich und Weizenecker im Jahre 2005 in [1] vorgestellt. Die hohe räumliche und zeitliche Auflösung wurde bereits in 2D-Phantom-Experimenten [2] als auch in 3D in-vivo-Experimenten [3] gezeigt. In den 3D-Experimenten wurde ein Bolus innerhalb eines schlagenden Mäuseherzens dargestellt.

Das physikalische Prinzip von MPI beruht auf der nichtlinearen Magnetisierungskurve von magnetischen Nanopartikeln. Angeregt durch ein oszillierendes sinusförmiges Anregungsfeld (Drive-Field), geraten die Nanopartikel periodisch in Sättigung, so dass die Magnetisierung der Partikel nun nicht mehr rein sinusförmig ist, sondern auch Harmonische der Anregungsfrequenz enthält. Anhand der Amplituden der Harmonischen kann die Menge der vorhandenen Partikel bestimmt werden.

Zur örtlichen Kodierung wird dem Drive-Field ein statisches Gradientenfeld (Selection-Field) überlagert. Die Feldstärke ist an einem Punkt – dem sogenannten feldfreien Punkt (FFP) – null. An allen anderen Orten hat sie durch einen steilen Magnetfeldgradienten betragsmäßig einen hohen Wert. An diesen Punkten sind die Nanopartikel stets in Sättigung, so dass nur Partikel in der Nähe des

FFPs auf das Anregungssignal reagieren. Durch Bewegung des FFPs durch den Betrachtungsbereich (field of view, FOV) kann ein Bild der Partikelverteilung generiert werden.

Der FFP kann entlang verschiedener Trajektorien durch das FOV gefahren werden. Als eine der besten Möglichkeiten haben sich Lissajous-Kurven herausgestellt [4]. Diese werden zurzeit auch in allen bekannten MPI-Systemen eingesetzt [2, 3, 5]. Die erzielte räumliche Auflösung ist umso besser, je dichter die Trajektorie ist. Jedoch verlängert sich die Aufnahmezeit auch mit der Dichte der Trajektorie. Ein Problem ist, dass die Frequenzen fest gewählt werden, da die Komponenten eines MPI-Systems auf diese fest eingestellt werden müssen. So werden zum Beispiel die Schwingkreise der Filter und der Sendespulen auf die jeweilige Resonanzfrequenz abgestimmt. Verschiedene Frequenzen auf einem Sendekanal sind nur mit sehr großem Aufwand realisierbar. Diese festgewählten Frequenzen resultierten bisher in einer festen örtlichen Auflösung und Aufnahmezeit. In diesem Beitrag wird eine neue Möglichkeit vorgestellt, die es erlaubt, die Aufnahmezeit und die örtliche Auflösung flexibel gegeneinander zu verändern. Hierzu wird die Phase einer Anregungsfrequenz so verändert, dass mit mehreren kurzen Trajektorien das gleiche Abtastverhalten erzielt wird wie mit einer langen und dichteren Trajektorie.

2 Material and Methoden

Eine 2D-Lissajous-Kurve entsteht durch Überlagerung zweier Wechselfelder mit leicht unterschiedlichen Frequenzen f_x und f_y . Das Frequenzverhältnis kann dabei so gewählt werden, dass

$$\frac{f_x}{f_y} = \frac{N}{N+1} \quad (1)$$

gilt. Je größer N gewählt wird, umso dichter wird die resultierende Trajektorie und umso länger wird auch die Repetitionszeit

$$T_r = \frac{N}{f_x} = \frac{N+1}{f_y} \quad (2)$$

Mit den Amplituden I_x und I_y und den Phasen φ_x und φ_y lassen sich die Ströme der einzelnen Richtungen mittels

$$\mathbf{I} = \begin{pmatrix} I_x \\ I_y \end{pmatrix} = \begin{pmatrix} \hat{I}_x \sin(2\pi f_x t + \varphi_x) \\ \hat{I}_y \sin(2\pi f_y t + \varphi_y) \end{pmatrix} \quad (3)$$

berechnen. Die Phasen φ_x und φ_y wurden dabei bislang fest gewählt. Statt die Dichte der Trajektorie über das Frequenzverhältnis (1) zu wählen, schlagen wir vor, bei einem ebenfalls festen Frequenzverhältnis mehrere Trajektorien mit unterschiedlichen Phasen φ_y aufzunehmen. Die Dichte der Gesamttrajektorie kann nun über die Anzahl P der verschiedenen Trajektorien geregelt werden. Dabei wird das Frequenzverhältnis so geändert, dass

$$\frac{f_x}{f_y} = \frac{N_{\text{kurz}}}{N_{\text{kurz}} + 1} = \frac{N}{N + P} \quad (4)$$

Tabelle 1. Daten der verwendeten Trajektorien.

N	N_{kurz}	P	f_x	f_y	T_r	T_{total}
99	99	1	25.000 Hz	25.252 Hz	$99 \cdot 0,4 \text{ ms} = 39,6 \text{ ms}$	$1 \cdot 3,6 \text{ ms} = 39,6 \text{ ms}$
99	9	11	25.000 Hz	27.777 Hz	$9 \cdot 0,4 \text{ ms} = 3,6 \text{ ms}$	$11 \cdot 3,6 \text{ ms} = 39,6 \text{ ms}$

gilt, wobei $N = N_{\text{kurz}} \cdot P$ ist. Die Phasen der P Einzeltrajektorien lassen sich mittels

$$\varphi_{y,p} = 2\pi \frac{p-1}{P} \quad \text{mit } p \in [1, P] \quad (5)$$

berechnen. Die Repetitionszeit T_{total} der Gesamtrajektorie ergibt sich aus

$$T_{\text{total}} = P \cdot T_r \quad (6)$$

In Tabelle 1 sind Daten für ein Beispiel mit $N = 99$, $N_{\text{kurz}} = 9$ und $P = 11$ gegeben. Die 11 Einzeltrajektorien sind in Abbildung 1 dargestellt und die Phasen dieser Trajektorien in Tabelle 2 gegeben. Vergleicht man die kurzen Trajektorien mit der langen in Abbildung 2(a) erkennt man deutlich, dass diese eine geringere Dichte aufweisen. Werden jedoch alle 11 Einzeltrajektorien zu einer Gesamtrajektorie überlagert, wie in Abbildung 2(b) dargestellt, hat diese die vergleichbare Dichte wie die lange Trajektorie.

Zur Evaluierung der verschiedenen Trajektorien, wurde der Kleintierscanner aus [3, 6] mittels [7] und [8] simuliert. Das genutzte Phantom (Abb. 4(a)) bestand aus unterschiedlich großen Zylindern (1,6 mm, 1,2 mm, 0,8 mm, 0,04 mm), die mit Nanopartikeln gefüllt sind. Als mathematisches Modell für die Partikel

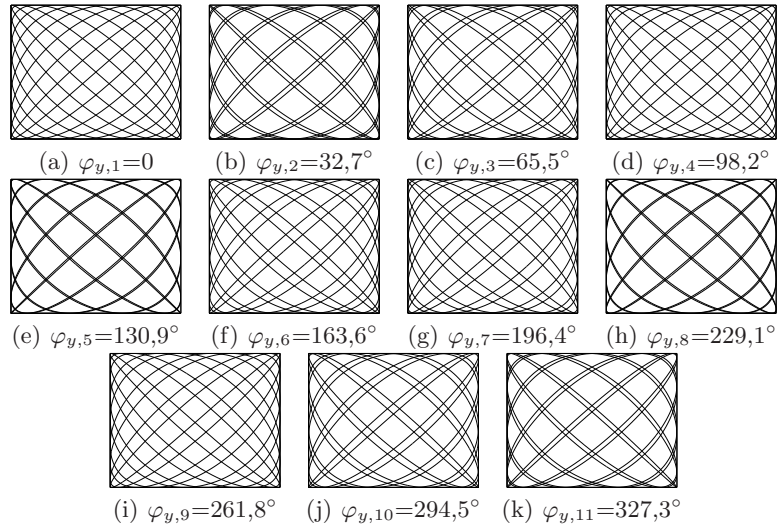
**Abb. 1.** Kurze Trajektorien mit unterschiedlichen Phasen φ_y .

Tabelle 2. Phasen der Einzeltrajektorien.

p	1	2	3	4	5	6	7	8	9	10	11
$\varphi_{y,p}$	0°	$32,7^\circ$	$65,5^\circ$	$98,2^\circ$	$130,9^\circ$	$163,6^\circ$	$196,4^\circ$	$229,1^\circ$	$261,8^\circ$	$294,5^\circ$	$327,3^\circ$

wurde die Langevin-Theorie des Paramagnetismus [9] verwendet. Hierzu wurde der Durchmesser des Eisenoxidkerns mit 30 nm angenommen.

3 Ergebnisse

Das Rekonstruktionsergebnis für die lange Trajektorie ist in Abbildung 4(b) dargestellt und für eine einzelne kurze Trajektorie mit $\varphi_y = 0^\circ$ in Abbildung 4(c). Es lässt sich erkennen, dass bei der langen Trajektorie wesentlich mehr Strukturen zu unterscheiden sind und somit eine höhere räumliche Auflösung erzielt wird als bei der kurzen Trajektorie. In Abbildung 4(d) ist das rekonstruierte Bild zu sehen, wenn die kurzen einzelnen Trajektorien mit unterschiedlicher Phase kombiniert werden. Die Rekonstruktion hat eine vergleichbare örtliche Auflösung, wie die der langen Trajektorien. Es ist somit möglich, die lange Trajektorie in verschiedene kurze Trajektorien aufzuteilen.

4 Diskussion

In diesem Beitrag wurde ein Verfahren für Magnetic Particle Imaging vorgestellt, das es ermöglicht, die feste Aufnahmezeit und Auflösung variabel zu verändern. Es wurde gezeigt, dass durch Variation der Phase bei einer Lissajous-Kurve verschiedene Trajektorien entstehen. Durch Kombination dieser Trajektorien ist es möglich, sowohl Aufnahmen mit einer sehr kurzen Repetitionszeit und niedriger Auflösung als auch Aufnahmen mit einer sehr hohen Auflösung und einer langen Repetitionszeit durchzuführen, ohne dass die Systemhardware verändert werden muss. Dies ermöglicht zum Beispiel eine schnelle Aufnahme zur Orientierung in

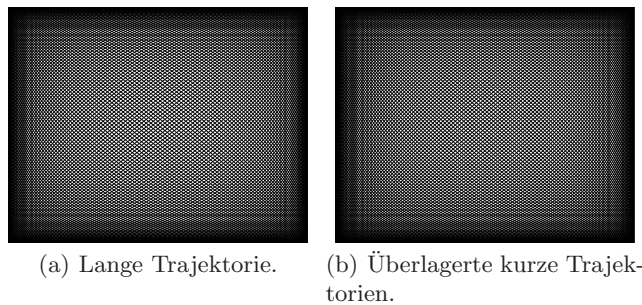
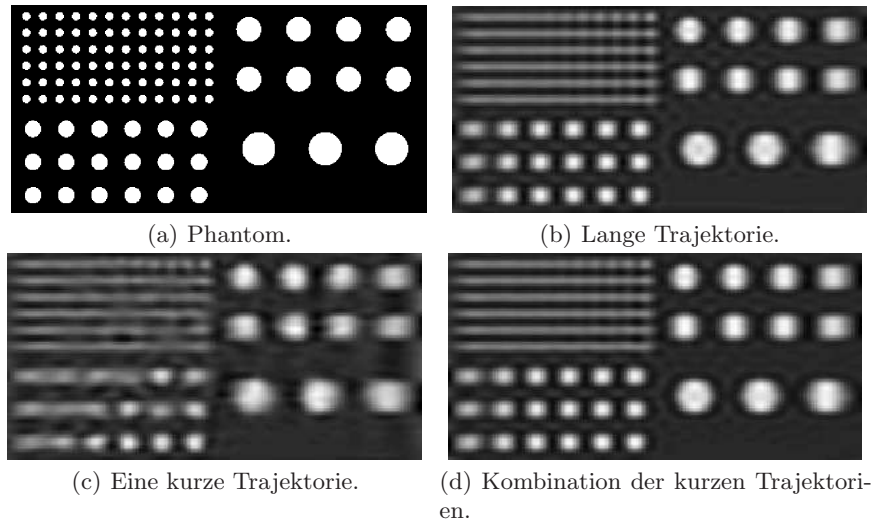
**Abb. 2.** Verschiedene Gesamt-Trajektorien.

Abb. 3. Rekonstruktionsergebnisse mit den verschiedenen Trajektorien.

niedriger Auflösung und anschließend eine hochaufgelöste Detailaufnahme. Weiterhin kann so bei einer dynamischen Bildgebung die Scanzeit angepasst werden. So kann zum Beispiel für die Aufnahmen des schlagenden Herzens eine niedrigere Aufnahmezeit gewählt werden als für Aufnahme von statischen Strukturen, die dafür in einer höheren Auflösung dargestellt werden können.

Literaturverzeichnis

1. Gleich B, Weizenecker J. Tomographic imaging using the nonlinear response of magnetic particles. *Nature*. 2005;435(7046):1214–7.
2. Gleich B, Weizenecker J, Borgert J. Experimental results on fast 2D-encoded magnetic particle imaging. *Phys Med Biol*. 2008;53(6):N81–N84.
3. Weizenecker J, Gleich B, Rahmer J, et al. Three-dimensional real-time in vivo magnetic particle imaging. *Phys Med Biol*. 2009;54(5):L1–L10.
4. Knopp T, Biederer S, Sattel T, et al. Trajectory analysis for magnetic particle imaging. *Phys Med Biol*. 2009;54(2):385–97.
5. Sattel T, Knopp T, Biederer S, et al. Single-sided device for magnetic particle imaging. *J Phys D Appl Phys*. 2009;42(1):1–5.
6. Schmale I, Gleich B, Kanzenbach J, et al. An introduction to the hardware of magnetic particle imaging. In: *Proc Med Phys Biomed Eng*. 25/II. Munich; 2009. p. 450–3.
7. Weizenecker J, Gleich B, Borgert J. Magnetic particle imaging using a field free line. *J Phys D Appl Phys*. 2008;41(10):3pp.
8. Rahmer J, Weizenecker J, Gleich B, et al. Signal encoding in magnetic particle imaging. *BMC Med Imaging*. 2009;9.
9. Chikazumi S, Charap S. *Physics of Magnetism*. New York: Wiley; 1964.

Time-of-Flight Kameras für die intraoperative Oberflächenerfassung

Alexander Seitel¹, Thiago R. dos Santos¹, Sven Mersmann¹, Jochen Penne²,
Ralf Tetzlaff^{1,3}, Hans-Peter Meinzer¹, Lena Maier-Hein¹

¹Abteilung für Medizinische und Biologische Informatik, DKFZ Heidelberg

²Lehrstuhl für Mustererkennung, Friedrich-Alexander Universität Erlangen

³Abteilung Radiologie, DKFZ Heidelberg

a.seitel@dkfz-heidelberg.de

Kurzfassung. Die Erfassung von Organoberflächen zur intraoperativen Registrierung spielt eine immer größere Rolle bei computergestützten Eingriffen. Während einige Techniken zur Organoberflächenerfassung (z.B. Laserscanning oder strukturiertes Licht) bereits zum Einsatz kamen, ist die Verwendung der neuartigen Time-of-Flight (ToF)-Technologie noch nicht untersucht worden. Ziel dieser Studie war die Bestimmung der Qualität von aus ToF-Distanzdaten generierten Organoberflächen. Hierfür wurden ToF-Bilddaten von 5 verschiedenen Schweineorganen aufgenommen. Hochaufgelöste Computertomographie (CT)-Daten vom selben Objekt dienten als Basis für Referenzoberflächen. Die mittlere Distanz zwischen den ToF-Oberflächen und den Referenzoberflächen lag nach einer ICP-basierten Registrierung zwischen 2.9 ± 2.1 mm (Lunge) und 5.7 ± 6.7 mm (Niere). Die Anwendung eines Kantenerhaltenden Glättungsfilters führte zu einer Verbesserung von bis zu 27.6%. Die Ergebnisse zeigen, dass die ToF-Technologie eine interessante Alternative für die intraoperative Oberflächenaufnahme ist.

1 Einleitung

Time-of-Flight (ToF)-Kameras gewinnen bei der schnellen dreidimensionalen Objektoberflächenerfassung immer mehr an Bedeutung. Das Prinzip, um Tiefendaten aus dem Kamerabild zu erhalten, beruht auf der Laufzeitmessung des Lichtes [1] und benötigt keine zusätzlichen Merkmale zur Tiefenextraktion. Ein moduliertes Lichtsignal im Infrarotbereich wird von der Kamera ausgesendet und vom beobachteten Objekt reflektiert. Mit Kenntnis der Lichtgeschwindigkeit c und der Modulationsfrequenz f kann die Distanz d , die ein Objektpunkt von der Kamera hat, berechnet werden aus dem Phasenunterschied ϕ zwischen ausgesendetem und reflektiertem Signal: $d = \frac{c}{2f} \cdot \frac{\phi}{2\pi}$. Anwendungen sind z.B. in der industriellen Bildverarbeitung, der Automobilindustrie, oder auch der Mensch-Maschine-Interaktion zu finden [1]. Erst kürzlich wurde diese Technologie auch im Umfeld der medizinischen Bildverarbeitung in verschiedenen Forschungsprojekten wie zur Atembewegungserfassung [2], zur Patientenpositionierung [3] sowie in einem neuartigen ToF-Endoskop [4] erprobt. Die Erfassung einer Organoberfläche für die intraoperative Registrierung entwickelt sich zu einer zentralen

Anforderung von Computerassistenzsystemen und endoskopischen Eingriffen [5]. Die ToF-Technologie könnte die hier gängigen jedoch deutlich komplexeren Verfahren wie Stereoendoskopie, Struktur aus Bewegung und strukturiertes Licht ersetzen. Der große Vorteil von ToF-Kameras ist, dass gleichzeitig ein Distanz-, Intensitäts- und Amplitudenwert für jedes Bildpixel mit einer hohen Wiederholungsrate (25 Aufnahmen pro Sekunde) zur Verfügung stehen. Jedoch sind diese Daten durch das komplizierte Messverfahren der Lichtlaufzeit mit Rauschen behaftet [6]. Deswegen ist es nötig, die Aufnahmen geeignet vorzuverarbeiten. In dieser Arbeit bestimmen wir die Qualität von Organoberflächen, die mit der ToF-Technologie erzeugt wurden. ToF-Aufnahmen wurden von fünf verschiedenen Schweineorganen erzeugt und mit Hilfe des kantenerhaltenden Bilateralfilters vorverarbeitet. Die Distanz zur Referenzoberfläche, gewonnen aus CT-Daten, diente als Gütemaß für die Qualitätsbestimmung.

2 Material und Methoden

Die *in-vitro* Evaluation erfolgte an explantierten Schweineorganen (Darm, Herz, Leber, Lunge, Niere) innerhalb einer Plastikbox (Abb. 3(a)). Von jedem Organ wurde ein CT-Bild mit 0.5 mm Schichtdicke und 0.3 mm Inkrement aufgenommen (Toshiba Aquilion 16 slice multidetector CT-scanner, Toshiba, Tokyo, Japan). Zusätzlich wurden 10 zeitlich direkt aufeinander folgende ToF-Datensätze für jedes Organ aufgenommen (Camcube 2.0, PMDTec, Modulationsfrequenz 20 MHz, Integrationszeit 1500 μs). Von jedem Organ wurde das durchschnittliche Distanz- und Amplitudenbild berechnet, um so das Messrauschen zu verringern. Unsere Methode, um die Qualität der erhaltenen ToF-Oberfläche zu ermitteln, bestand aus folgenden Schritten:

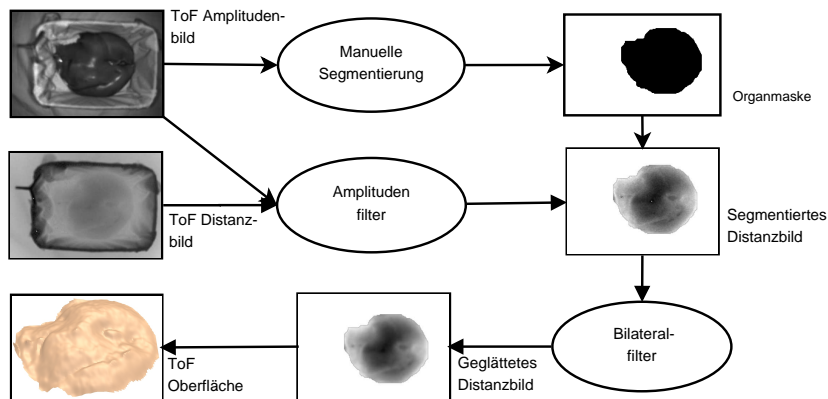


Abb. 1. Ablauf der Datenvorverarbeitung. Der Amplitudenfilter wurde benutzt, um Ausreißer ausserhalb des Amplitudenbereichs [20 %,80 %] [1] zu eliminieren. Das segmentierte Distanzbild wurde mit dem kantenerhaltenden Bilateralfilter geglättet.

- *Vorverarbeitung*: Um die Organoberflächen aus den CT-Daten zu erhalten, wurden die CT-Schichtbilder manuell segmentiert. Die ToF-Oberflächenerzeugung geschah nach dem Vorgehen beschrieben in Abb. 1. Mit Hilfe des Amplitudenbildes wurden die Distanzpixel eliminiert, deren Amplitudenwert ausserhalb des Amplitudenbereiches von [20 %, 80 %] liegen. Das resultierende Distanzbild wurde mithilfe einer Organmaske, erhalten aus einer manuellen Organsegmentierung des Amplitudenbildes, segmentiert. Der Bilateralfilter (mit empirisch bestimmten Parametern $\sigma_d = 1$ und $\sigma_r = 80$), welcher die Glättung sowohl unter Berücksichtigung der Pixelnachbarschaft als auch des Pixelwertes (in diesem Fall der Distanz) durchführt, wurde auf das segmentierte Distanzbild angewendet, um Tiefenmerkmale zu erhalten. Die Oberfläche wurde schließlich als gleichmäßiges Dreiecksnetz unter Berücksichtigung der Kameraparameter (Brennweite, Pixelabstand) aus dem geglätteten Bild berechnet.
- *Registrierung*: Zur Registrierung des Koordinatensystems der ToF-Kamera mit dem Koordinatensystem des CTs kam der Iterative Closest Point (ICP) Algorithmus zum Einsatz (Abb. 2). Initialisiert wurde er mit einer rigiden Transformation berechnet aus einer landmarkenbasierten Registrierung, wobei die Ecken der Box als Landmarken dienen.
- *Qualitätsbestimmung*: Um die Qualität der ToF-Oberfläche zu bestimmen, wurde die Distanz zur CT-Oberfläche als mittlere Distanz aller Abstände der Gitterpunkte der ToF-Oberfläche zum nächstgelegenen Punkt der CT-Oberfläche bestimmt.

3 Ergebnisse

Abb. 3(b) zeigt die Abstände zwischen den CT-Oberflächen und den geglätteten/originalen ToF-Oberflächen. Das beste Ergebnis bezüglich der geglätteten Oberfläche ist bei der Lunge mit einem mittleren Abstand von 2.9 ± 2.1 mm (original: 4.0 ± 3.4 mm) zu beobachten. Weiter folgen die Leber (geglättet: 3.0 ± 2.2 mm,

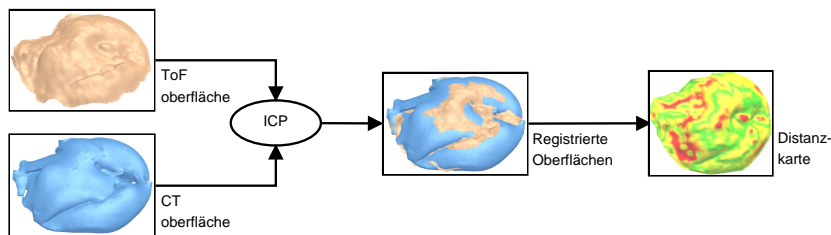


Abb. 2. Qualitätsbestimmung und Registrierung der ToF- und CT-Oberflächen. Die Registrierung wurde mit Hilfe des ICP Algorithmus durchgeführt. Die Abstandsmessung zwischen den registrierten Oberflächen wurde durch die Berechnung der Distanzen zwischen jedem Gitterpunkt der ToF-Oberfläche und dem entsprechend nächsten Punkt auf der CT-Oberfläche durchgeführt.

original: 3.3 ± 2.4 mm), das Herz (geglättet: 3.7 ± 2.6 mm, original: 4.1 ± 3.1 mm), der Darm (geglättet: 4.0 ± 2.9 mm, original: 4.4 ± 3.4 mm) und die Niere (geglättet: 5.7 ± 6.7 mm, original: 10.9 ± 6.7 mm). Die größte Verbesserung durch den Vorverarbeitungsansatz konnte bei der Lunge mit 27.6 % erzielt werden, gefolgt von der Niere mit 14.1 %, dem Herz mit 10.0 %, der Leber mit 9.7 % und dem Darm mit 8.2 %. Abb. 3(c) zeigt die originalen und geglätteten ToF-Oberflächen farbkodiert mit der Distanz zur Referenzoberfläche des CTs.

4 Diskussion

Unseres Wissens nach ist diese Studie die erste, welche die Qualität von Organoberflächen, aufgenommen mit einer ToF-Kamera, untersucht. Abhängig vom Organ konnte die Anwendung einer einfachen Vorverarbeitung den Fehler in der Distanz auf Werte zwischen 3 und 8 mm senken. Trotz dieser vielversprechenden Ergebnisse enthalten die erzeugten Oberflächen nach wie vor Bereiche in denen

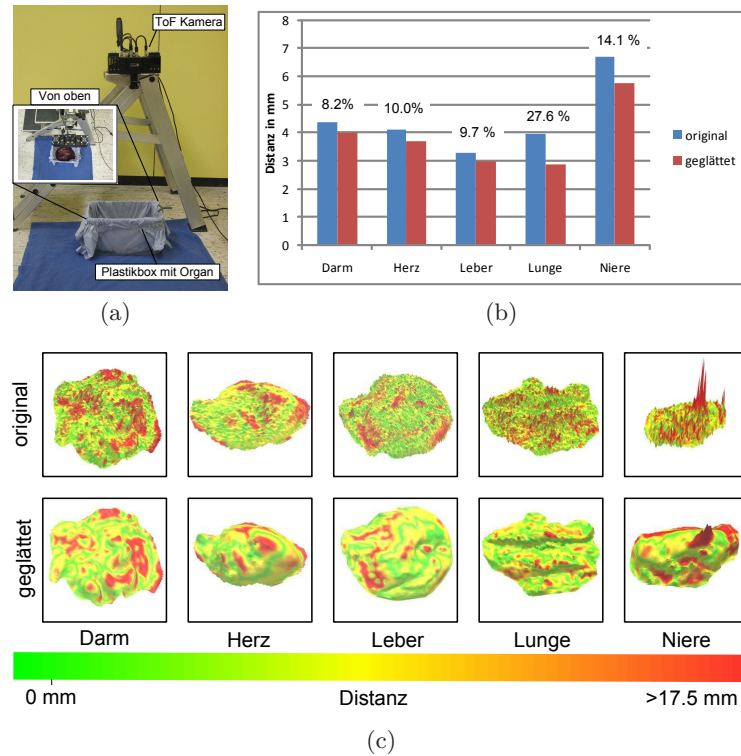


Abb. 3. (a) Experimenteller Aufbau bei der *in-vitro* Evaluation. (b) Mittlere Distanz zwischen der CT-Oberfläche und der originalen/geglätteten ToF-Oberfläche sowie die Verbesserung durch den Glättungsansatz in %. (c) Original (obere Reihe) und geglättete (untere Reihe) Oberflächen der einzelnen Organe, farbkodiert mit den Distanzwerten.

der Distanzfehler weit über 1 cm lag Abb. 3(c). Es scheint, dass diese Abweichungen hauptsächlich auf oder in der Nähe von herausstehenden Oberflächenstrukturen auftraten, wo zusätzlich eine hohe Lichtreflexion zu beobachten war. Das kann zum einen damit erklärt werden, dass die Integrationszeit für diese Fälle zu hoch angesetzt wurde, was an diesen Stellen zu übersättigten Pixeln und somit zu unvorhersagbarem Rauschen führte. Zum anderen könnte auch der intensitätsabhängige Distanzfehler sowie der Störeinfluss der Reflexion eine Rolle spielen. Eine anwendungsspezifische Optimierung der Kameraparameter (Integrationszeit, Modulationsfrequenz) sowie des Amplitudenfilters (z.B. dynamische Anpassung des erlaubten Amplitudenbereichs) könnten die ToF-Oberflächen weiter verbessern. Weiterhin muss untersucht werden, inwieweit der Bilateralfilter optimal für die Rauschunterdrückung bei ToF-Oberflächen ist bzw. wie eine adaptive Abwandlung die Oberflächenergebnisse weiter verbessern könnte. Zusätzlich sollten weitere Vorverarbeitungsschritte wie Kamerakalibrierung zur Minderung des Distanzfehlers (intensitätsabhängig, hervorgerufen durch Aliasing oder Umgebunglicht) oder andere angepasste Filteransätze (z.B. Diffusionsfilterung) das Ergebnis weiter verbessern. Schließlich müssen weitere Untersuchungen zeigen, ob die vorgestellten Methoden für die Anwendung der intraoperativen Organregistrierung geeignet sind und wie gut die ToF-basierte Oberflächenaufnahme im Vergleich zu anderen Verfahren wie strukturiertem Licht oder Laserscanning ist. Nichtsdestotrotz zeigen die Ergebnisse, dass die relativ neue ToF-Technologie eine interessante Alternative zur intraoperativen Oberflächenaufnahme ist.

Danksagung. Die vorliegende Studie wurde im Rahmen des DFG-geförderten Graduiertenkolleg 1126 „Intelligente Chirurgie“ durchgeführt. Dank geht an Martina Jochim von der Abteilung Radiologie im DKFZ für die Unterstützung während der CT-Aufnahmen.

Literaturverzeichnis

1. Kolb A, et al. Time-of-Flight sensors in computer graphics. In: Eurographics. CH-1288 Aire-la-Ville: Eurographics; 2009. p. 119–34.
2. Schaller C, et al. Time-of-flight sensor for respiratory motion gating. *Med Phys.* 2008;35(7):3090–3.
3. Schaller C, et al. Time-of-flight sensor for patient positioning. *Proc SPIE.* 2009;7261:10–1–8.
4. Penne J, et al. Time-of-flight 3-D endoscopy. *Lect Notes Computer Sci.* 2009;5761:467–74.
5. Cash DM, et al. Concepts and preliminary data toward the realization of image-guided liver surgery. *J Gastrointest Surg.* 2007;11:844–59.
6. Frank M, et al. Theoretical and experimental error analysis of continuous-wave time-of-flight range cameras. *Opt Eng.* 2009;48(1):013602–1–16.

Laparoscopic Quantitative 3D Endoscopy for Image Guided Surgery

Jochen Penne¹, Christian Schaller¹, Rainer Engelbrecht³, Lena Maier-Hein²,
Bernd Schmauss³, Hans-Peter Meinzer², Joachim Hornegger¹

¹Pattern Recognition Lab and Erlangen Graduate School in Advanced Optical Technologies (SAOT), Friedrich-Alexander-University Erlangen-Nuremberg, Germany

²Division of Biological and Medical Informatics, German Cancer Research Center, Heidelberg, Germany

³Chair of Microwave Engineering and High Frequency Technology,
Friedrich-Alexander-University Erlangen-Nuremberg, Germany
Jochen.Penne@informatik.uni-erlangen.de

Abstract. Image-guided surgery aims at relating data of an on-going intervention to preoperative information about important anatomical structures or tumor localizations. Using intra-operatively acquired 3D information of the operation area has been proven to be the right starting point for accomplishing this task, but is leading to technical challenges especially in the field of minimally invasive surgery. A novel laparoscopic 3D surface scan device providing >3000 3D surface points at 20 fps was utilized to exemplarily accomplish a) the fusion of 3D surface scans and color information at 20 fps and b) the fusion of 3D surface scans and CT data with a mean error of 1.68 mm. Experiments were on one hand accomplished in-vitro using a porcine stomach and on the other hand using a liver phantom. The results verify the proof of concept of the utilized system and lead to clear directions for future research.

1 Introduction

Segmental liver resections as well as ablative therapies require accurate and precise tumor localization. In the preoperative setting the anatomic location of hepatic tumors is commonly determined using tomographic imaging techniques, computed tomography (CT) and magnetic resonance imaging (MRI). Transferring this information to the intra-operative setting poses a technical challenge as the peritoneal cavity is subject to deformations caused by respiration as well as topologic changes of the operation site during the intervention. Therefore, mapping the intra-operative position of utilized instruments onto preoperative, high-resolution, tomographic data sets, will enhance the tumor localization and resection.

As the intra-abdominal application of traditional frameless stereotaxy approaches is facing few stationary landmarks and considerable intra-operative deformation, current research has successfully focused on acquiring 3D surface

scans of the operation area even in minimally invasive interventions [1]. The registration of these 3D surface scans to preoperative tomographic data relates the on-going resection to the tumor and surrounding anatomical structures. The registration has to take into account organ deformations. The accuracy of locating intra-parenchymal targets is comparable for traditional fiducial-based registrations and the rather new surface-based registrations [2].

The contribution of this paper is two-fold: a) A novel 3D surface scanning device for minimally invasive interventions is described and b) initial results for the registration of the 3D surface scans with color information and CT data are provided.

2 Materials and Methods

For acquiring 3D surface scans a rather novel scanning device was utilized. The used Time-of-Flight-based (ToF based) 10 mm laparoscope acquires 3072 3D surface points at 20 fps with an average precision of 0.89 mm [3]. The operation area is scanned with a spatial and temporal resolution which is comparable if not superior to alternative techniques like laser range scanners while only utilizing off-the-shelf technology components.



Fig. 1. The utilized endoscope system: To the upper laparoscope a conventional color sensor (768×576 px) was proximally mounted; to the lower laparoscope a ToF sensor for acquiring 3D surface scans and gray-value images was proximally mounted. By actively illuminating the operation area via the laparoscope with an optical reference signal and the pixel-wise measurement of the phase-delay between emitted and received reference signal in a proximally mounted ToF sensor the length of the propagation path of the light is estimated in each pixel of the 64×48 pixel (model 3 kS from PMDTechnologies GmbH) ToF sensor matrix [4]. Additionally, a gray-value image of the field-of-view is provided by the ToF sensor. Depth map and gray-value image provided by the ToF sensor are registered by construction.

After accomplishing a calibration of the projective properties [5] of the endoscope optic 3D Cartesian coordinates can be computed from the measured depth maps [6]. In order to express these 3D coordinates in a Euclidean coordinate system whose origin coincides with the endoscope tip and whose z-axis is aligned with the optical axis of the endoscope the measured distances have to be expressed relative to the endoscope tip. This requires an additional distance calibration step. In contrast to the distance calibration proposed in [3] a more extensive distance calibration which also takes into account systematic integration-time related distance measurement errors of the ToF sensor [7] was accomplished.

For correcting lens distortions in the depth map it is advisable to compute the inverse distortion correction numerically [8] instead of warping the image and interpolate undistorted pixel values.

For the experiments a second laparoscope was mounted fixed to the ToF laparoscope and the relative transformation between the two optics was determined by a calibration routine, which provides the extrinsic and intrinsic camera parameters. A color image sensor was proximally mounted to the additional laparoscope. By projecting the 3D points of the surface onto the image plane of the additional laparoscope using the computed relative transformation the fusion of color and ToF data was accomplished. During the projection lens distortions were taken into account, too.

Two experiments were accomplished using the described dual-laparoscope device:

1. A manually insufflated porcine stomach was inspected in order to acquire in-vitro data and accomplish the fusion of ToF laparoscope and color laparoscope data.
2. A liver phantom was inspected. The 3D surface scans were registered with the liver surface phantom data available from a CT image volume using the iterative closest point (ICP) algorithm [9].

3 Results

Fig. 2 shows the checker-board overlay of the fusion of the gray-value image of the ToF sensor and the color image which was computed by projecting the 3D surface scan onto the color image plane. Additionally, a 3D surface scan textured with the registered color information is displayed. The data was acquired during experiment 1 (inspection of the porcine stomach). On a dual-core CPU (2.4 GHz, 2 GB RAM) the computation of the fusion took 47 ms, where all computations

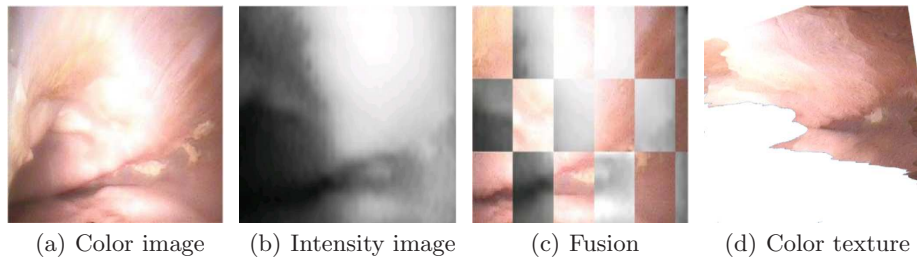


Fig. 2. Results of porcine stomach in-vitro experiments. Two endoscope optics were utilized: One with a proximally mounted color sensor and one with a proximally mounted ToF sensor for 3D surface scans. An occlusion detection was omitted. (a): color data; (b): gray-value image which is provided in addition to the 3D surface scan from the ToF sensor; (c): checker-board overlay of registered color and 3D surface scan data; (d): 3D surface scan data which was textured by the registered color data.

regarding the ToF data were accomplished on an increased lateral resolution of 320×240 pixels compared to the original 64×48 pixels of the ToF sensors. Increasing the lateral resolution by bicubical interpolation leads to a denser sampling of the color image when projecting the ToF 3D points onto the color image plane. Fig. 3 shows the fusion of the laparoscopic 3D surface scans of the liver phantom (experiment 2) with the 3D point cloud extracted from the CT data. As the 3D surface scan was registered before with color information, it is textured with this color information. Computing the average value of the closest distance between a 3D point from the ToF data and its nearest neighbor from the 3D point cloud of the CT data yielded a value of 1.68 mm.

4 Discussion

The obviously well overlay of gray-value and color data in Fig. 2 verifies that the calibration of the projective properties as well as the distance calibration lead to a sufficient measurement accuracy as well as a correct 3D modeling of the scene: Any error in the estimation of the projective properties or the distance

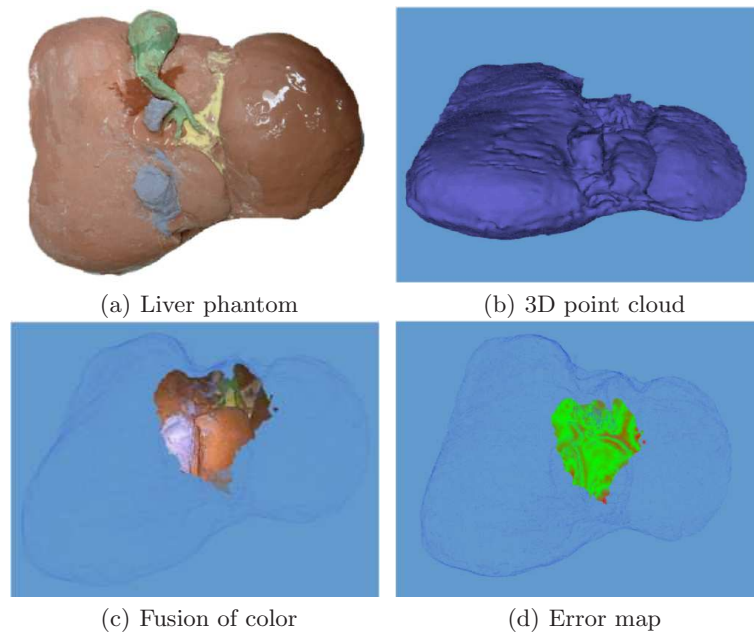


Fig. 3. Results of multimodal fusion. (a): utilized liver phantom; (b): 3D point cloud which was extracted from a CT data set; (c): fusion of 3D surface scan with color and the 3D CT data set point cloud; (d): color-encoded distance error between a 3D point of the 3D surface scan and its closest point from the CT 3D point cloud (bright/green: 0.04 mm, dark/red: 7.65 mm; average error: 1.68 mm).

calibration would lead to a mismatch between color and ToF laparoscope and a misalignment.

The fusion of 3D surface scans and CT data yields a mean point-to-closest-point Euclidean distance error of 1.68 mm. The initial ICP transformation was computed based on nine manually selected corresponding landmarks in CT and 3D surface scans. This can be accomplished intra-operatively under sterile conditions, for example, by using a ToF based gesture control as proposed in [10] and is no issue decreasing the practical feasibility of the approach. The authors of this paper envision a system where this intra-operative registration is done once and the future fusion of 3D surface scans and CT data is based on the tracking of the endoscope optics.

The presented results are preliminary: Experiments involving surface fiducials are subject to current research and will lead to an additional validation of the fusion of 3D surface scans and color information as well as 3D surface scans and CT data. The last mentioned data fusion does currently not involve any deformation compensation: This is the focus of future work.

References

1. Rauth TP, Bao PQ, Galloway RL, et al. Laparoscopic surface scanning and subsurface targeting: Implications for image-guided laparoscopic liver surgery. *Surgery*. 2007;142(2):207–14.
2. Cash DM, Sinha TK, Chapman WC, et al. Incorporation of a laser range scanner into image-guided liver surgery: surface acquisition, registration, and tracking. *Med Phys*. 2003;30(7):1671–82.
3. Penne J, Höller K, Stürmer M, et al. Time-of-Flight 3D endoscopy. *Lect Notes Comput Sci*. 2009;5761:467–74.
4. Xu Z, Schwarte R, Heinol H, et al. Smart pixel: photometric mixer device (PMD). In: *Proc Int Conf Mechatron Mach Vis Pract*; 1998. p. 259–64.
5. Zhang Z. A Flexible New Technique For Camera Calibration. *IEEE Trans Pattern Anal Mach Intell*. 2000;22(11):1330–4.
6. Linarth AG, Penne J, Liu B, et al. Fast fusion of range and video sensor data. In: Valldorf J, Gessner W, editors. *Proc Adv Microsys Automot Appl*. Berlin; 2007. p. 119–34.
7. Kolb A. ToF Sensors in Computer Graphics. In: *Proc EuroGraphics*; 2009. p. 259–264.
8. de Villiers J, Leuschner F, Geldenhuys R. Centi-pixel accurate real-time inverse distortion correction. In: *Proc Int Symp Optomechatron Techn*. vol. 7266; 2008. p. 1–8.
9. Kynor DB, Friets EM, Knaus DA, et al. Tissue localization using endoscopic laser projection for image-guided surgery. *Proc SPIE*. 2005;5744:225–35.
10. Soutschek S, Penne J, Hornegger J, et al. 3D gesture-based scene navigation in medical imaging applications using time-of-flight cameras. In: *Omnipress*, editor. *Proc IEEE CVPR*; 2008.

Partikelfilterung für die Kompensation von Atembewegung während der navigierten Bronchoskopie

Ingmar Gergel¹, Thiago R. dos Santos¹, Ralf Tetzlaff², Lena Maier-Hein¹,
Hans-Peter Meinzer¹, Ingmar Wegner¹

¹Abt. Medizinische und Biologische Informatik, Deutsches Krebsforschungszentrum

²Abt. Radiologie, Deutsches Krebsforschungszentrum

`i.gergel@dkfz-heidelberg.de`

Kurzfassung. Obwohl viele Forschungsgruppen auf dem Bereich der navigierten Bronchoskopie arbeiten, ist das Problem der stabilen Positionsbestimmung noch nicht gelöst, da es sich als schwierig erweist die Veränderung der Lunge durch die Atembewegung mit zu berücksichtigen. Ein in der Literatur beschriebener Ansatz basiert auf Partikelfilterung. Die Evaluation erfolgte jedoch ohne Berücksichtigung der Atembewegung. Ziel dieser Arbeit ist es, Partikelfilterung ausführlich unter Atembewegung zu evaluieren. Des Weiteren soll untersucht werden inwieweit eine Parametrisierung des Partikelfilters in Abhängigkeit des Bronchusdurchmessers sinnvoll ist. Hierfür wurde eine Simulationsumgebung basierend auf einem aus Human CT-Datensätzen berechneten Deformationsfeld verwendet. Den Ergebnissen zufolge schneidet die Partikelfilterung besser als vorherige Verfahren ab und ist zur Kompensation der Lungenbewegung während einer navigierten Bronchoskopie geeignet.

1 Einleitung

Mit einem Bronchoskop kann das Innere der Lunge durch einen optischen Kanal untersucht werden. Je tiefer die zu untersuchende Region liegt, desto schwieriger ist es ein über das Bronchoskop eingeführte Instrument an eine definierte Stelle zu bringen. Mehrere Forschungsgruppen [1, 2, 3] arbeiten an elektromagnetischen Navigationssystemen zur Unterstützung von diagnostischen und therapeutischen Bronchoskopien. Allen gemein ist das Ziel die aktuelle Position eines Instruments relativ zu einem präinterventionell aufgenommenen, statischen CT anzuzeigen. Die wesentliche Herausforderung bei der Positionsbestimmung des Instruments ist es, die Veränderung der Lunge durch die Atembewegung zu berücksichtigen.

Wegner et al. beschreiben ein Verfahren basierend auf inkrementeller Echtzeitregistrierung. Die Position eines getrackten Bronchoskops wird hierbei auf die Zentrallinie des Tracheobronchialbaums mittels Nächste-Nachbarn-Algorithmus Centerline Matching (CM) korrigiert [3]. Somit stellen Sie sicher, dass die gefilterte Position sich immer im Innern der Bronchien befindet. Allerdings wird auch beschrieben, dass das Verfahren dazu neigt die Instrumentenposition kurzzeitig in einem falschen Nachbarbronchus anzuzeigen.

Atmosukarto et al. verwenden ebenfalls ein Verfahren basierend auf inkrementeller Echtzeitregistrierung jedoch mit Schwerpunkt auf Partikelfilterung. Hierbei wird die zeitliche Abfolge vorheriger Positionen während der Bronchoskopie mitberücksichtigt, um so eine möglichst präzise Positionsbestimmung zu erzielen. Die Evaluation erfolgte aber ausschließlich an einer rigiden, getrockneten Schaflunge [4] und berücksichtigt somit keine Lungenbewegung.

Ziel dieser Arbeit ist es, mit einer überarbeiteten Version der in [5] vorgestellten Simulationsumgebung, Partikelfilterung ausführlich unter Atembewegung zu evaluieren, um so zu untersuchen ob es ein geeignetes Verfahren zur Positionsbestimmung während einer bronchoskopischen Interventionen ist.

2 Methoden

Zur Evaluation der Partikelfilterung unter Atembewegung diente die in [5] vorgestellte Simulationsumgebung als Grundlage. Sie ermöglicht es, eine hohe Anzahl an Interventionen unter Berücksichtigung der wichtigsten Störeinflüsse in einer klinischen Umgebung zu simulieren. Es werden die Ungenauigkeit des Tracking-systems (elektromagnetische Rauschen), die Verzerrungen aufgrund metallischer Gegenstände (OP-Tisch) und die Atembewegung simuliert.

Da der ursprünglichen Atembewegungssimulation jedoch stark vereinfachte Annahmen zugrunde liegen, wurde ein neues Deformationsmodell basierend auf Human CT-Datensätzen erstellt. Hierfür wurden zwei Thorax-CT Aufnahmen in Inspirations- und Expirationslage verwendet. Anhand manuell bestimmter Korrespondenzen konnte über eine Elastic-Body-Spline-Transformation die Deformation der Lunge neu simuliert werden.

Mit dem in die Simulationsumgebung neu integrierten Deformationsmodell wurden zuerst die Parameter für den Partikelfilter auf einem Trainingsdatensatz eingestellt. Als Partikelfilter wurde eine Erweiterung des in [6] beschriebenen Algorithmus verwendet. Der Filter wurde dahingehend modifiziert, dass alle Partikel ausschließlich auf der Zentrallinie des Tracheobronchialbaums verteilt werden (Centerline Particle Filter, CPF). Für die Optimierung der CPF-Parameter n (Anzahl an Partikel) und r (Radius in dem die Partikel verteilt werden) wurde eine Kostenfunktion f_0 angewandt. f_0 wurde hierbei definiert als der prozentuale Anteil an richtig klassifizierten Positionen. Eine Position gilt als richtig klassifiziert, wenn sie sich im selben Gefäß wie die zugehörige Referenzposition (Ground Truth) befindet.

Der Parameter r wurde zusätzlich zum einen (a) als Konstante und zum andern (b) als variabler Wert in Abhängigkeit des Bronchusdurchmesser (r_B) untersucht. Für (b) wurden beide Möglichkeiten analysiert: r kleiner und größer als der momentane Bronchusdurchmesser r_B .

Auf einem zweiten Testdatensatz erfolgte anschließend die eigentliche Evaluation des CPF und der Vergleich zu reinem CM. Um zu untersuchen wie viele Positionen im richtigen Bronchus angezeigt werden, wurde p_T als der prozentuale Anteil an richtig klassifizierten Positionen berechnet. p_T wurde ebenfalls für die Gefäßgenerationen p_{0-8} (p_0 = Trachea, p_8 = tiefster automatisch zu segmen-

tierender Bronchus) einzeln berechnet. Zum besseren Vergleich von p_{0-8} wurde p_{ratio} als prozentualer Anteil an Positionen in der jeweiligen Gefäßgeneration ermittelt.

Des Weiteren wurden zwei Maße reg_T und irr_{max} zur Quantifizierung der Gleichmäßigkeit von CPF und CM berechnet: je kleiner der Wert desto gleichmäßiger der Algorithmus. Hierbei soll erfasst werden, wie kontinuierlich der jeweilige Algorithmus über die Zeit ist, beziehungsweise inwieweit ein sprunghafter Wechsel der Position von einem Bronchus in den Nachbarbronchus vorliegt. reg_T wurde definiert als mittlerer quadratischer Abstand von der gefilterten Position zu dem Zeitpunkt t und $t-1$. Das Maß irr_{max} beschreibt die mittlere maximale Abweichung als $\frac{\sum d_m}{k}$ mit $d_m = \max |p_t - p_{t-1}|$ pro Intervention, p = gefilterte Position zum Zeitpunkt t bzw. $t-1$ und k = Anzahl simulierter Interventionen. Unabhängig von der Gleichmäßigkeit, wurde zusätzlich rms_T als der mittlere quadratische Abstand der gefilterten Position zur Referenzposition berechnet, solange richtig klassifiziert wurde.

Abschließend wurde in einem zweiten Versuchsaufbau das Deformationsmodell dahingehend verändert, dass es die Bewegung einer Hochfrequenzbeatmung (HFV) simuliert. Unter HFV versteht man eine künstliche Beatmung welche häufig während bronchoskopischen Interventionen verwendet wird, da sie bei sehr hohen Beatmungsraten (>65 Beatmungen/Minute) sehr kleine Atemzugvolumen ermöglicht und somit eine geringe Lungenbewegung erzeugt.

3 Ergebnisse

Insgesamt wurden über 1000 Interventionen sowohl auf dem Trainingsdatensatz als auch auf dem Testdatensatz simuliert und untersucht. Anhand des Trainingsdatensatzes wurden über die Kostenfunktion f_0 die Parameter $n = 150$ und $r = 7$ mm empirisch bestimmt. Ein variabler Radius r in Abhängigkeit des Bronchusdurchmesser zeigte keine Verbesserung im Vergleich zu einem über alle Bronchien hinweg konstant gewählten Radius.

Abbildung. 1 zeigt eine simulierte Intervention nach CPF (a) und CM (b). Der gefilterte Pfad nach CM weist mehrere Lücken auf. Ebenso werden einige Positionen in einem falschen Bronchus angezeigt. Der Pfad nach CPF ist hingegen kontinuierlich und durch eine weniger sprunghafte Positionsbestimmung gekennzeichnet.

Quantitativ wird dieses Verhalten in p_T bestätigt (Tab. 1). Auf dem Testdatensatz wurden nach CPF durchschnittlich über 91 % der Positionen richtig klassifiziert (94 % auf dem Trainingsdatensatz), bei CM hingegen lediglich 86 % (89 % auf dem Trainingsdatensatz). Die beiden Maße zur Bestimmung der Gleichmäßigkeit der Algorithmen reg_T und irr_{max} sind in beiden klinischen Szenarien nach CPF um mehr als die Hälfte kleiner als nach CM, (Tab. 2). Bei der reinen Abstandsbestimmung rms_T der gefilterten Position zur Referenzposition weist das Verfahren nach CM kleinere Werte auf.

Tabelle 1. Testdaten in % an richtig klassifizierten geschätzten Positionen.

Generation	spontane Atmung		HFV		p_{ratio}
	CPF	CM	CPF	CM	
p_0	93.3	87.2	93.3	93.0	11.4
p_1	93.8	73.8	90.0	90.1	25.2
p_2	90.5	90.8	93.9	93.8	21.7
p_3	89.3	84.4	90.4	92.4	14.7
p_4	92.0	78.0	93.3	79.6	10.2
p_5	59.5	67.8	68.5	78.3	8.2
p_6	54.0	49.9	88.9	71.4	6.2
p_7	56.5	73.3	66.4	82.7	1.2
p_8	92.0	86.7	100	100	0.7
p_T	91.6 ± 7.4	86.0 ± 7.6	96.2 ± 3.2	94.1 ± 3.4	

4 Diskussion

In der vorliegenden Arbeit wurde CPF als ein Algorithmus vorgestellt, welcher die bereits zurückgelegte Strecke berücksichtigt, um so während einer navigierten Bronchoskopie die Atembewegung zu kompensieren. Für die Evaluation des Ansatzes unter realitätsnahen Bedingungen wurde ein neues Lungendeformati-

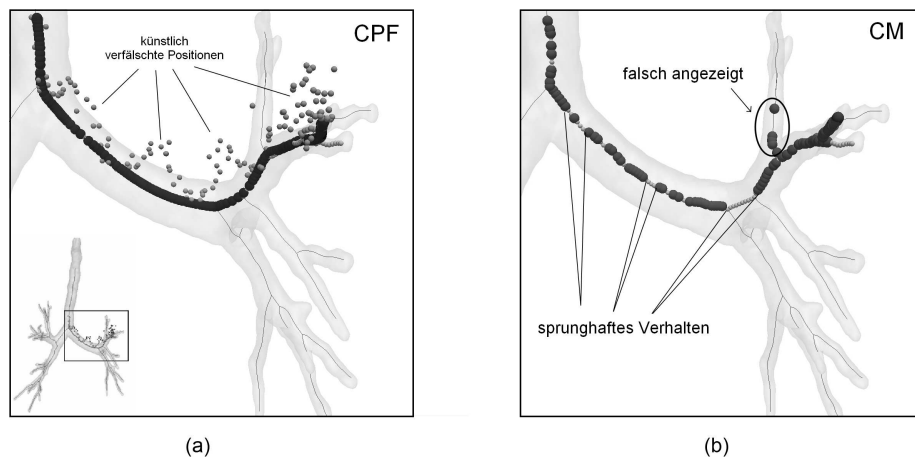


Abb. 1. Eine simulierte Intervention unter spontaner Atmung mit CPF und CM. Die hellen Punkte repräsentieren in (a) die künstlich verfälschten Positionen aufgrund der Atmung und EM-Rauschen; die dunklen Punkte zeigen den kontinuierlichen Pfad nach CPF. In (b) repräsentieren die hellen Punkte den Referenzpfad; die dunkleren Punkte zeigen den lückenhaften Pfad nach CM.

Tabelle 2. Gleichmäßigkeit beider Algorithmen (alle Maße in mm).

	spontane Atmung			HFV		
	reg _T	irr _{max}	rms _T	reg _T	irr _{max}	rms _T
CPF	1.1	5.6	4.2 ± 1.8	0.9	1.9	5.9 ± 1.6
CM	2.9	12.7	1.4 ± 1	2.0	5.9	3.0 ± 2.2

onsfeld entwickelt und damit die Parameter des CPF optimiert. Eine vom Bronchusdurchmesser abhängige Parameterisierung zeigte keine Vorteile gegenüber einem konstanten Radius. Abschließend wurde CPF bezüglich der Genauigkeit bei der Positionsbestimmung insgesamt, als auch in jeder Gefäßgeneration einzeln, evaluiert.

Solange richtig klassifiziert wurde, kann auf den ersten Blick vermutet werden, dass CM aufgrund kleinerer rms_T Werte eine genauere Positionsbestimmung liefert. Die Maße reg_T und irr_{max} zeigen jedoch, dass CPF insgesamt in einem robusteren und gleichmäßigeren Pfad resultiert (Abb. 1). Ferner zeigt insbesondere p_T, dass während CM die Position häufiger im falschen Bronchus angezeigt wird, als es bei CPF der Fall ist: 86 % vs. 91 %.

Zusammenfassend lässt sich sagen, dass den Ergebnissen zufolge Centerline Particle Filtering das bereits bestehende Centerline Matching übertrifft und zur Kompensation der Atembewegung während einer elektromagnetisch navigierten bronchoskopischen Interventionen geeignet ist.

Literaturverzeichnis

1. Schwarz Q, Mehta A, Ernst A, et al. Elektromagnetic navigation during flexible bronchoscopy. *Respiration*. 2003; p. 516–22.
2. Mori K, Deguchi D, Akiyama K, et al. Hybrid bronchoscope tracking using a magnetic tracking sensor and image registration. In: *Proc MICCAI*. vol. 2; 2005. p. 543–50.
3. Wegner I, Biederer J, Tetzlaff R, et al. Evaluation and extension of a navigation system for bronchoscopy inside human lungs. *Proc SPIE*. 2007.
4. Atmosukarto I, Soper TD, Glenny RW, et al. An interactive 3D user interface for guided bronchoscopy. *Proc SPIE*. 2007.
5. Wegner I, Tetzlaff R, Biederer J, et al. An evaluation environment for respiratory motion compensation in navigated bronchoscopy. *Proc SPIE*. 2008;6918.
6. Isard M, Blake A. Condensation: conditional density propagation for visual tracking. *Int J Computer Vis*. 1998;29(1):5–28.

Sight-based Magnification System for Surgical Applications

Anabel Martin-Gonzalez¹, Sandro M. Heining², Nassir Navab¹

¹Computer Aided Medical Procedures, TU München

²Trauma Surgery Department, LMU

`martingo@in.tum.de`

Abstract. This work presents the development of an augmented reality magnification system implemented in a head-mounted display for surgical applications. The system provides a magnified view with a novel control based on tracked sight orientation. The system was evaluated by measuring the completion time of a suturing task performed by surgeons. The magnifying approach implemented provided global context of the operating field. The sight-based activation was widely accepted by surgeons as a useful functionality to control viewing modalities.

1 Introduction

Augmented reality (AR) is defined as a technology in which the user's view of the real world is enhanced or augmented with additional information generated by a computer [1]. In the medical domain, AR technology has been used to assist surgical procedures, including the development of medical virtual tools [2].

Medical AR systems based on the use of surgical microscopes have been developed and tested. Birkfellner et al. [3] introduced an adapted commercial head-mounted operating microscope for stereoscopic augmented reality visualization. In a more recent work, Garcia et al. [4] developed an AR system to align preoperative tomographic images into a surgical microscope with a rigidly mounted mini-tracking system to track movements of surgical tools and the patient.

Despite high quality of real optics in head-mounted microscope AR systems, the continuous magnified view aligned to the eyes' sight, disables the surgeon to visualize other areas in the operating room (OR) with normal vision. In addition, the linear magnification occludes visual information below the magnified space, restricting the observer to perceive complete information surrounding the operating field. In order to overcome these limitations, we propose a head-mounted magnification system with sight-based activation.

2 Material and Methods

2.1 Hardware Setup

The hardware consists of a video-see through head-mounted display (HMD) with two color cameras attached to the HMD (Fig. 1), a pointer tool, and two infrared based tracking systems to have high-precision tracking.

The internal tracking system, termed RAMP [5], consist of a single infrared camera, mounted on the HMD, for tracking a reference frame to provide the HMD pose with high rotational accuracy.

The external tracking system (A.R.T. GmbH) consists of a set of four infrared cameras attached to the ceiling that covers a working area of approximated 2.5m^3 . The target's tracking accuracy is around 0.35mm . Since this is not a single, but multiple camera system, it requires an arrangement with a minimum of four fiducial markers for obtaining the six degrees of freedom (DOF) position of the target. This system is used for tracking the pointer tool's tip, the reference frame and the HMD pose in case that the reference frame target is out of scope of the RAMP system's infrared camera.

2.2 Sight-based Activation

Our system provides an intuitive control to activate and deactivate the magnified view by means of tracking the HMD's sight. When the user focuses on the operating field the magnification will be turned on. Focusing outside this area it will be turned off, allowing the surgeon the freedom to perform other tasks with normal vision, such as grabbing surgical instruments, monitoring medical devices in the OR, or interacting visually with medical partners.

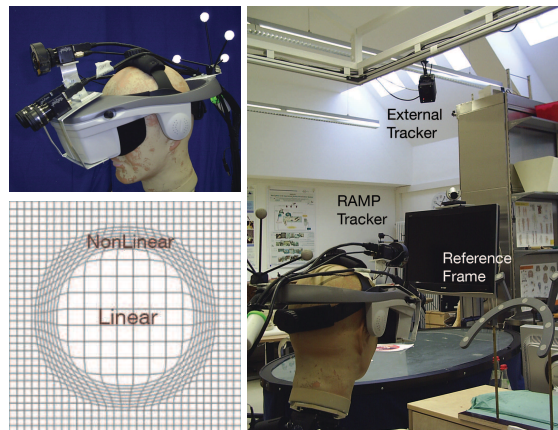


Fig. 1. Head-mounted magnification system (top-left). Hybrid magnification approach (bottom-left). Tracking system setup (right).

In order to establish when the user is focusing on the operative field, a vector of sight and its distance from the working area are determined as follows. First, the position of the pointer’s tip is saved to define location of operating field. Then, the distance d between the saved location and left camera of the HMD is calculated. Then the angle α between the camera’s sight vector s and the normalized distance vector \hat{d} is computed based on the dot product of two vectors, i.e., $\alpha = \cos^{-1}(\hat{d} \cdot s)$. Experimentally, we defined that if $d < 50$ cm and $\alpha < 15^\circ$, then the magnification is activated. Camera location and orientation were obtained from the translation vector and rotation matrix contained in the following transformation:

$$\text{ext}T_{\text{cam}} = \text{ext}T_{\text{ref}} \cdot (\text{ramp}T_{\text{ref}})^{-1} \cdot (\text{cam}T_{\text{ramp}})^{-1} \quad (1)$$

where $\text{ext}T_{\text{cam}}$ denotes transformation from external tracking system to HMD’s left camera, $\text{ext}T_{\text{ref}}$ is from external tracking system to reference frame, $\text{ramp}T_{\text{ref}}$ is from RAMP system to reference frame and $\text{cam}T_{\text{ramp}}$ is from left camera to RAMP system.

2.3 Hybrid Magnification Approach

Different methods for implementing general non-linear magnification transformations are presented by Keahey and Robertson [6]. We have adapted and applied these concepts into an HMD for AR.

The hybrid magnification approach implemented consists of a radial linear magnification surrounded by a radial non-linear magnification with constrained domain (Fig. 1). The non-linear magnification transforms each point in the domain as follows: given a center point of magnification $C = (C_x, C_y)$ and a point to transform $P = (P_x, P_y)$, let $\hat{P} = P - C$. Find the radius component of the polar coordinates of \hat{P} so that $r = \sqrt{\hat{P}_x^2 + \hat{P}_y^2}$. The new coordinates are then $C + \frac{h(r)}{r}\hat{P}$, where $h(r) = \tanh(r)$.

The linear magnification presents a uniform level of magnification across the domain. In a constrained magnification, transformations are performed inside a sub-area of the domain; all points outside the domain remain untransformed. This preserves global context.

3 Results

Five surgeons from our clinical partner, experts on microsurgery, were volunteers for performing a suturing procedure assisted by our system (Fig. 3). The experiment consisted on practicing two single knots on a suturing module for testing two different magnification modes ($2\times$ factor): hybrid and linear (Fig. 2).

First, the artificial wound location is determined, by means of the tracked pointer, and saved in the system. Then, the pointer is removed and the surgeon starts suturing with a randomly selected mode. This is repeated with the

Table 1. Suturing task completion time. Linear & hybrid mode were set to 3 & 2 and 2 & 3 for first and second trial, respectively. All measures are given in [ms].

	All trials		First trial		Second trial	
	Linear	Hybrid	Linear	Hybrid	Linear	Hybrid
Average	161.40	157.40	171.67	203.00	146.00	127.00
Standard deviation	24.17	43.31	26.31	9.90	12.73	15.39

remaining mode on a second suturing module. The results of experiments are presented in Tab. 1.

Due to the learning curve intrinsic in the experiments' repetition and adaptation to the AR system, we were expecting that the time for performing the second trial would be faster than the first one. Thus, we consider relevant to analyze first and second trials independently. Three surgeons were evaluated first with linear mode and the remaining two with hybrid mode. In contrast, the second trial had two surgeons with linear and three with hybrid.

It is very interesting to notice from these results that, during the first contact with the AR system (first trial), the surgeon performed the suturing procedure faster with linear mode (magnification similar to the medical loupe, with which the surgeon is completely familiarized). On the other hand, after the user was

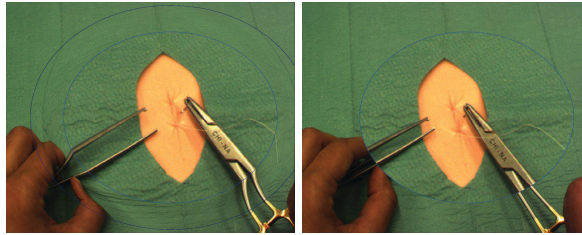


Fig. 2. Magnification modes displayed into the HMD: hybrid (left) and linear (right).



Fig. 3. Surgeon assisted by: the magnification system (left) and a surgical loupe attached to the glasses (right).

adapted to the AR system (second trial), it took less time for suturing with hybrid mode.

4 Discussion

A magnification system with sight-based activation was developed for surgical applications. The sight-based activation control supplied a novel functionality. The magnification approach provides continuity between magnified and non-magnified areas, prevents occlusions and provides a global context of the operating scene.

By wearing a surgical loupe (Fig. 3), the user can also switch between regular and magnified views by moving the eyes; however, the user will get a semi-obstructed and restricted view. The magnification system allows to automatically recovering a complete field of view. Moreover, an AR system could provide additional features taking advantages of digital manipulation which optical devices could never offer.

As advances in hardware will provide the scientific community with higher resolution and lighter HMDs, the novel software solution, such as the one presented here, pave the path towards integration of AR into medical applications.

Acknowledgement. This work was granted partly by Bayerische Forschungsförderung (BFS) and Secretaría de Educación Pública de México (SEP).

References

1. Koller D, Klinker G, Rose E, et al. Real-time vision-based camera tracking for augmented reality applications. In: VRST-97; 1997. p. 87–94.
2. Navab N, Feuerstein M, Bichlmeier C. Laparoscopic virtual mirror: new interaction paradigm for monitor-based augmented reality. In: Virtual Reality Conference, IEEE. USA; 2007. p. 43–50.
3. Birkfellner W, Figl M, Huber K, et al. A head-mounted operating binocular for augmented reality visualization in medicine: design and initial evaluation. *IEEE Trans Med Imaging*. 2002;21:991–7.
4. Garcia J, Caversaccio M, Pappas I, et al. Design and clinical evaluation of an image-guided surgical microscope with an integrated tracking system. *Int J CARS*. 2007;1:253–64.
5. Sauer F, Khamene A, Vogt S. An augmented reality navigation system with a single-camera tracker: System design and needle biopsy phantom trial. In: Proc MICCAI; 2002. p. 116–24.
6. Keahey TA, Robertson EL. Techniques for Non-Linear Magnification Transformations. In: Proc IEEE Visualization; 1996. p. 38–45.

Automatische Landmarkendetektion und -übertragung zur Evaluation der Registrierung von thorakalen CT-Daten

René Werner, Jan-Christoph Wolf, Jan Ehrhardt,
Alexander Schmidt-Richberg, Heinz Handels

Institut für Medizinische Informatik, Universitätsklinikum Hamburg-Eppendorf
r.werner@uke.uni-hamburg.de

Kurzfassung. Es wird ein Verfahren zur automatischen Detektion und Übertragung von Landmarken in 4D Lungen-CTs präsentiert. Die Landmarken können z.B. zur Evaluation von Registrierungsverfahren eingesetzt werden. Um charakteristische Punkte der Lunge als Landmarkenkandidaten zu ermitteln, wird ein krümmungsbasierter Differentialoperator genutzt. Weitere Anforderungen an die Detektion wie eine gleichmäßige Verteilung der Landmarken in der Lunge werden berücksichtigt. Zur Landmarkenübertragung wird ein Template-Matching-Verfahren eingesetzt. Die Robustheit der Übertragung wird durch kombinierte Nutzung von Intensitätsinformationen und Differentialeigenschaften der Bilddaten gewährleistet. Eine Evaluation des Verfahrens anhand von fünf 4D-CT-Daten zeigt, dass vorwiegend anatomisch relevante Gefäßverzweigungen detektiert und mit hoher Genauigkeit übertragen werden.

1 Einleitung

Der klinische Einsatz von Registrierungsverfahren bedarf einer verlässlichen Abschätzung der Registrierungsgenauigkeit. Diese ist im Kontext von nicht-linearer Registrierung schwierig, da die Ground-Truth-Deformation zumeist nicht bekannt ist [1]. In diesem Beitrag wird die nicht-lineare Intrapatienten-Registrierung von 4D Lungen-CT-Daten adressiert. Die Registrierungsgenauigkeit wird i.d.R. mittels lungeninterner Landmarken validiert, die in den zu registrierenden Daten lokalisiert und zur Berechnung des Target-Registration-Errors (TRE) herangezogen werden [1, 2]. Eine manuelle Lokalisation insbesondere größerer Landmarkensets ist zeitintensiv und fehleranfällig. Ein Verfahren zur Unterstützung der Registrierungsvalidierung sollte folglich idealerweise vollständig automatisiert sein (Anforderung A1), ausreichend viele Landmarkenkandidaten detektieren, die (verifizierbar!) charakteristische Punkte der Lunge beschreiben (A2) und gleichmäßig in der Lunge verteilt sind (A3). Die Landmarken sollten verlässlich zwischen den unterschiedlichen CT-Daten des betrachteten Patienten übertragen werden (A4). Charakteristische Punkte der Lunge, wie z.B. Gefäßverzweigungen, zeichnen sich durch spezielle Krümmungseigenschaften aus und können deshalb durch Differentialoperatoren detektiert werden [3]. In [4] werden solche Punkte semi-automatisch mit haptischer Unterstützung detektiert. Allerdings können

nicht ausreichend viele Landmarken bestimmt werden. In [1] wird ein Verfahren zur Detektion ausreichend vieler und gleichmäßig verteilter Landmarken beschrieben. Allerdings ist die Übertragung semi-automatisch und es werden keine Krümmungseigenschaften berücksichtigt. In diesem Beitrag wird ein automatisches Verfahren, das die genannten Anforderungen erfüllt, beschrieben, evaluiert und exemplarisch zur Validierung eines Registrierungsansatzes eingesetzt.

2 Material und Methoden

2.1 Methodische Beschreibung des Verfahrens

Das Verfahren zur Landmarkendetektion ist in Algorithmus 1 skizziert. Die Suchregion wird anhand einer Maske auf die Lunge beschränkt, die durch eine atlasbasierte Segmentierung ermittelt wird [5]. Um charakteristische Punkte gem. (A2) zu finden, wird ein sogenannter Distinctiveness-Wert berechnet [1]. In [1] wurde ein Term gewählt, der auf dem Betrag des Intensitätsgradienten basiert. Um typische Landmarken wie Bifurkationen von Bronchial- und Gefäßbaum zu detektieren, erscheint der Einsatz krümmungsbasierter Operatoren zielgerichteter [4]. Deshalb wird hier alternativ ein Distinctiveness-Term mittels des Förstner-Operators definiert. Die gleichmäßige Verteilung der Landmarken in der Lunge gem. (A3) wird gewährleistet, indem bei Auswahl der Landmarken eine minimale euklidische Distanz vorgegeben wird, die paarweise zwischen den Landmarken einzuhalten ist [1].

Die detektierten Landmarken werden mittels Template-Matching auf andere CT-Daten des Patienten übertragen (Templategröße 11^3 Voxel, zentriert um die Landmarkenposition; Suchregion im Zielbild: $5 \times 5 \times 10 \text{ cm}^3$, zentriert um die originäre Landmarkenposition). Es werden zwei Matching-Durchgänge pro Landmarke durchgeführt: zunächst auf Basis der CT-Daten bzw. Hounsfield-Units, dann unter ausschließlicher Berücksichtigung der Antwort des auf die CT-Daten angewendeten Förstner-Operators. Als gesuchte Landmarkenposition wird das Zentrumsvoxel der Zielbildregion gewählt, für die jeweilige Werte maximal mit jenen des Templates korrelieren. Um eine robuste Übertragung zu gewährleisten, werden Landmarkenkandidaten verworfen, für die Korrelationswerte von < 0.8 für den ersten Matching-Vorgang resultieren oder die beiden Matching-Durchgänge zu unterschiedlichen Zielregionen führen. Es werden dann alternative Landmarkenkandidaten gesucht, d.h. Detektion und Übertragung gekoppelt.

2.2 Validierung des Verfahrens: Durchgeführte Experimente

Die Validierung basiert auf CT-Daten von fünf Patienten, aufgezeichnet zu maximaler Ein- und Ausatmung (Auflösung: $1 \times 1 \times 1.5 \text{ mm}^3$). Anforderungen (A1) und (A3) sind durch Alg. 1 erfüllt. Zur Evaluation von (A2) werden in den CT-Daten zur Einatmung für die beiden beschriebenen Distinctiveness-Terme 100 Landmarken detektiert und von drei Experten auf ihre Eignung als „charakteristischer Punkt der Lunge“ geprüft. Den Experten ist nicht bekannt, welches Verfahren eingesetzt wurde („blinde“ Evaluation).

Algorithmus 1 Landmarkdetektion und -transfer.

Input: CT data $CT_{1,2} : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}$, lung segmentation $Seg_1 : \Omega \rightarrow \{0, 1\}$ Number n of the landmarks sought**Output:** List $\langle l_1 \rangle$ of landmarks in CT_1 ; list $\langle l_2 \rangle$ of corresponding landmarks in CT_2

- 1: Let $\mathbf{p}, \mathbf{p}' \in \Omega$ be the lung voxels in CT_1 , i.e. all voxels with $Seg_1 = 1$.
Then, compute for each \mathbf{p} the associated *distinctiveness* $D(\mathbf{p})$:

$$D(\mathbf{p}) := \frac{[\det \underline{C}(\mathbf{p}) / \text{trace } \underline{C}(\mathbf{p})]}{\max_{\mathbf{p}'} [\det \underline{C}(\mathbf{p}') / \text{trace } \underline{C}(\mathbf{p}')]}} \sum_{\mathbf{q} \in \mathbf{Q} \subset \mathbb{S}_r^2(\mathbf{p})} \frac{\text{MSD}(B_{r'}(\mathbf{p}), B_{r'}(\mathbf{q}))}{|Q|}$$

where $[\det \underline{C}(\mathbf{p}) / \text{trace } \underline{C}(\mathbf{p})]$ denotes the *Förstner-Operator* [3] and $\underline{C} := \overline{\nabla CT_1 (\nabla CT_1)^T}$ with ∇CT_1 the image gradient of CT_1 , $S_r^2(\mathbf{p})$: 2-sphere of radius r , centered at \mathbf{p} , $B_{r'}(\mathbf{p})$: 3-ball of radius r' , centered at \mathbf{p} .Let $(\mathbf{p}_1, \dots, \mathbf{p}_{|B_{r'}(\mathbf{p}) \cap \Omega|})$ and $(\mathbf{q}_1, \dots, \mathbf{q}_{|B_{r'}(\mathbf{q}) \cap \Omega|})$ be correspondingly sampled sequences of the voxels in $B_{r'}(\mathbf{p})$ and $B_{r'}(\mathbf{q})$. Then, MSD is defined as

$$\text{MSD}(B_{r'}(\mathbf{p}), B_{r'}(\mathbf{q})) := \frac{1}{|B_{r'}(\mathbf{p}) \cap \Omega|} \sum_{i=1}^{|B_{r'}(\mathbf{p}) \cap \Omega|} (CT_1(\mathbf{p}_i) - CT_1(\mathbf{q}_i))^2.$$

- 2: Sort points \mathbf{p} with $D(\mathbf{p}) \geq \theta_{D(\mathbf{p})}$ in descending order acc. to $D(\mathbf{p})$ into list $\langle p \rangle$.
 - 3: If size of $\langle p \rangle$ is $< n$, decrease distinctiveness threshold $\theta_{D(\mathbf{p})}$.
 - 4: If size of $\langle p \rangle$ is $< n$ and $\theta_{D(\mathbf{p})} > 0$, go to line 2.
 - 5: **for all** points \mathbf{p} in list $\langle p \rangle$ **do**
 - 6: Move \mathbf{p} to $\langle l_1 \rangle$ if its Euclidian distance to all points in $\langle l_1 \rangle$ is $> \theta_{\text{dist}}$.
 - 7: If size of $\langle l_1 \rangle$ is n , then continue with line 9.
 - 8: **end for**
 - 9: If size of $\langle l_1 \rangle$ is $< n$, decrease the minimum distance θ_{dist} .
 - 10: If size of $\langle l_1 \rangle$ is $< n$ and $\theta_{\text{dist}} > 0$, go to line 5.
 - 11: **for all** elements \mathbf{l}_1 in $\langle l_1 \rangle$ **do**
 - 12: Extract a $m_x \times m_y \times m_z$ subimage $T(\mathbf{l}_1)$ of CT_1 (the template to be matched), centered at \mathbf{l}_1 , and search the voxel $\mathbf{l}_2 \in \Omega$ such that the $m_x \times m_y \times m_z$ subimage $T'(\mathbf{l}_2)$ of CT_2 maximizes the correlation r_{HU} of the intensity values of T and T' .
 - 13: Analogously do a *template matching* for the Förstner images of CT_1 and CT_2 .
 - 14: **if** the correlation r_{HU} is < 0.8 **or** the returned voxels \mathbf{l}_2 of the template matching processes (lines 12 and 13) differ **then**
 - 15: Remove \mathbf{l}_1 from list $\langle l_1 \rangle$ and set $D(\mathbf{l}_1) = 0$.
 - 16: **else**
 - 17: Append voxel \mathbf{l}_2 to list $\langle l_2 \rangle$.
 - 18: **end if**
 - 19: **end for**
 - 20: If size of $\langle l_2 \rangle$ is $< n$, set $D = 0$ for all elements of $\langle l_1 \rangle$.
 - 21: If size of $\langle l_2 \rangle$ is $< n$ and there exist lung voxels with $D > 0$, go to line 2.
-

Um die Landmarkenübertragung zu evaluieren, werden den Experten dann für die Landmarken, die mittels des Förstner-Operators detektiert sind, die übertragenen Positionen präsentiert und beurteilt. Aus vorherigen Studien liegen für jeden Patienten zudem 80 zu beiden Atemphasen manuell lokalisierte Landmarken vor. Diese werden ebenfalls automatisiert übertragen und die Distanz zwischen übertragener und manuell bestimmter Landmarke ermittelt.

2.3 Anwendung zur Validierung der Registrierungsgenauigkeit

Das Verfahren wird exemplarisch zur Evaluation eines nicht-linearen Registrierungsverfahrens eingesetzt (nicht-parametrisch, elastische Regularisierung, Thirion-Kraftterme [6]). Es werden verschiedene Landmarkensets betrachtet (50, 100, 300 Landmarken). Für die geschätzten Bewegungsfelder zwischen Ein- und Ausatmung wird der TRE berechnet. Als Vergleich liegen TRE-Werte vor, die aus den manuell lokalisierten Landmarken ermittelt wurden.

3 Ergebnisse

72% der Landmarken, die mittels des auf dem Förstner-Operator basierenden Distinctiveness-Terms detektiert wurden, wurden als „charakteristischer Punkt der Lunge“ beurteilt (Abb. 1). Bei 21% waren sich die Experten nicht sicher, bei 7% konnten sie die Charakteristik des Punktes nicht erkennen. Analoge Werte für das Kriterium gem. [1] sind 60%, 25% und 15%.

Korrelationsschwellwert und Kombination der Intensitäts- und Krümmungsinformationen führten dazu, dass sämtliche betrachtete Landmarken als erfolgreich übertragen eingeschätzt wurden. Für ein Template-Matching nur auf Basis

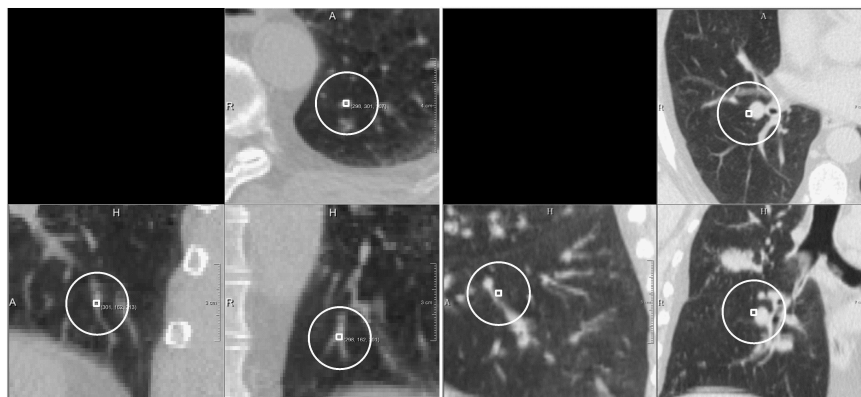


Abb. 1. Zum Begriff von „charakteristischen Punkten innerhalb der Lunge“. Links: Die Position beschreibt eine Bifurkation des Gefäßbaumes und somit einen günstigen Landmarkenkandidaten (Detektion mittels Förstner-Operators). Der rechts abgebildete Landmarkenkandidat (Distinctiveness auf Basis des Intensitäts-Gradientenbetrags) ist schwer reproduzierbar identifizierbar, d.h. ein ungünstiger Kandidat.

der Hounsfield-Units resultierten hingegen Fehlerraten von bis zu 7%. Die Gegenüberstellung von automatischer Übertragung und manueller Landmarkendetektion ergab Abweichungen von 1.2 ± 0.1 mm (Vergleich: Intraobserver-Variabilität der manuellen Landmarkendetektion = 0.9 ± 0.1 mm).

Für die automatisch detektierten Landmarkensets ergaben sich mittlere TRE-Werte von 0.7 ± 0.1 mm (50 Landmarken), 0.8 ± 0.1 mm (100 Landmarken) und 1.0 ± 0.1 mm (300 Landmarken), d.h. eine Registrierungsgenauigkeit unterhalb der Voxelgröße. Korrespondierende Landmarken-Abstände vor Registrierung waren 7.2 ± 1.5 mm, 7.4 ± 1.4 mm und 7.6 ± 1.6 mm. Die Übereinstimmung mit den Werten aus der manuellen Landmarkendetektion (TRE: 1.1 ± 0.1 mm; vor Registrierung: 7.4 ± 2.0 mm) belegen die Verlässlichkeit des Verfahrens.

4 Diskussion

Es wurde ein robustes automatisches Verfahren zur Landmarkendetektion und -übertragung in (4D-)Lungen-CT-Daten präsentiert. Der Einsatz eines krümmungsbasierten Differentialoperators zur Detektion von charakteristischen Punkten der Lunge hat sich als gewinnbringend erwiesen. Es zeigt sich jedoch, dass die Landmarkendetektion noch weiter optimiert werden kann und sollte. Die Übertragung der Landmarken zwischen den CT-Daten eines Patienten mittels Template-Matching ist für die betrachteten Daten ausgesprochen verlässlich. Die somit ermöglichte Detektion und Übertragung von (auch großen) Landmarkensets ohne Interaktionsaufwand erleichtert es, umfassendere Evaluationsstudien von z.B. Verfahren zur Registrierung von Lungen-CT-Daten durchzuführen und so detailliertere Einsichten in deren Verhalten und Funktion zu erhalten.

Danksagung. Die präsentierten Projektbeiträge wurden von der Deutschen Forschungsgemeinschaft gefördert (DFG, HA 2355/9-1).

Literaturverzeichnis

1. Murphy K, van Ginneken B, Pluim JPW, et al. Semi-automatic reference standard construction for quantitative evaluation of lung CT registration. Proc MICCAI. 2008;11:1006–13.
2. Sarrut D, Delhay B, Villard PF, et al. A comparison framework for breathing motion estimation methods from 4-D imaging. IEEE Trans Med Imaging. 2007;26:1636–48.
3. Hartkens T, Rohr K, Stiehl HS. Evaluation of 3D operators for the detection of anatomical point landmarks in MR and CT images. Comput Vis Image Underst. 2002;86:118–36.
4. Färber M, Gawenda B, Bohn CA, et al. Haptic landmark positioning and automatic landmark transfer in 4D Lung CT Data. Proc BVM. 2008; p. 313–7.
5. van Rikxoort EM, Prokop M, de Hoop B, et al. Automatic segmentation of the pulmonary lobes from fissures, airways, and lung borders: Evaluation of robustness against missing data. Proc MICCAI. 2009; p. 263–71.
6. Modersitzki J. Numerical Methods for Image Registration. Oxford Univ. Press; 2003.

Graphbasierte Registrierung von Tubulären Strukturen

Thiago R. dos Santos, Ingmar Gergel, Sven Mersmann, Hans-Peter Meinzer,
Lena Maier-Hein

Abteilung für Medizinische und Biologische Informatik, DKFZ, Heidelberg
t.santos@dkfz.de

Kurzfassung. Verschiedene Ansätze zur Registrierung medizinischer Bilddaten basieren auf der Extraktion und Registrierung baumartiger Strukturen (z.B. Blutgefäße, Bronchialbäume). In diesem Beitrag präsentieren wir eine Methode zur schnellen Korrespondenzsuche in anatomischen Bäumen, die keine initiale Positionierung erfordert. Basierend auf einer Graphrepräsentation der Bäume wird die Ähnlichkeit jeweils zweier Knoten (Verzweigungspunkte) sowohl anhand topologischer Eigenschaften als auch anhand anatomischer Merkmale bewertet. Die Korrespondenzermittlung erfolgt anschließend durch die globale Maximierung einer Ähnlichkeitsmatrix. Den Ergebnissen einer Evaluationsstudie an Schweinebronchien zufolge liefert die hier vorgestellte Methode eine hohe Rate korrekter Zuordnungen im Vergleich zu verwandten Arbeiten.

1 Einleitung

Ansätze zur Registrierung medizinischer Bilddaten basieren zunehmend auf der Extraktion und Registrierung baumartiger Strukturen wie den Bronchien oder den Blutgefäßen [1]. Die Bestimmung von Korrespondenzen gestaltet sich durch verrauschte Eingangsdaten, Organbewegung und Artefakte in diesem Zusammenhang jedoch schwierig. Ein von Metzen *et al.* [1] dargelegter Überblick über vorhandene Verfahren zur Korrespondenzsuche zeigt, dass graphbasierte Verfahren in diesem Kontext an Bedeutung gewinnen. Graham and Higgins verwenden einen Algorithmus [2, 3], welcher zwei Baumstrukturen anhand zweier Deformationsmodelle auf einen Referenzbaum abbildet. Allerdings fehlt eine umfassende Evaluation. Zudem verwenden die Autoren die Parallelität zwischen Kanten als Ähnlichkeitsmaß, wodurch sich eine Abhängigkeit von der relativen Lage der Bäume zueinander ergibt. Die Arbeit von Metzen *et al.* [1] ist unabhängig von der Lage der Bäume, basiert jedoch auf einer “Maximum Clique” Suche¹, welche bekannterweise ein NP-vollständiges Problem darstellt und nur lokal optimale Lösungen liefert (alle Zuweisungen sind auf eine bestimmte Region begrenzt).

Unserer Kenntnis nach wurde noch kein Verfahren vorgestellt welches (1) schnell, (2) unabhängig von der relativen Lage der Bäume zueinander sowie (3)

¹ Eine Clique in einem Graph ist ein Teilgraph, in welchem alle Knoten paarweise durch Kanten verbunden sind.

robust gegenüber Deformationen ist und (4) eine global verteilte Knotenzuweisung ermöglicht. In dieser Arbeit stellen wir ein Verfahren vor, welches diesen Forderungen genügt.

2 Methoden

Der Ablauf für die Registrierung tubulärer Strukturen durch graphbasierte Korrespondenzsuche ist in Abbildung 1 am Beispiel von Bronchialbäumen dargestellt. Zunächst werden aus den segmentierten Bronchien die Zentrallinien extrahiert [4]. Basierend auf einer Graphdarstellung der Strukturen, in der jeder Knoten (außer der Wurzel) mit einer Verzweigung im entsprechenden Bronchialbaums korrespondiert, wird anschließend eine Korrespondenzsuche durchgeführt (Abschnitt 2.1), um jedem Knoten eines Baumes maximal einen Knoten des anderen Baumes zuzuweisen. Nach einer Reduktion von falschen Zuweisungen (Abschnitt 2.2) kann mit Hilfe der gefundenen Korrespondenzen eine Transformation zur Registrierung der Eingangsbilder berechnet werden.

2.1 Korrespondenzsuche

Unser Ansatz zur Korrespondenzsuche basiert auf der Berechnung einer Ähnlichkeitsmatrix zwischen den Knoten beider Bäume. Diese Matrix wird verwendet, um eine Knotenzuordnung zu finden, bei der die globale Ähnlichkeit maximal ist. Das Verfahren besteht aus drei Arbeitsschritten: *topologische Ähnlichkeitsauswertung*, *Merkmals-Ähnlichkeitsauswertung* und *Zuweisung*. Die Ähnlichkeitsmatrix besteht aus der Summe der Bewertungen der topologischen Ähnlichkeit und der Merkmals-Ähnlichkeit.

Topologische Ähnlichkeitsauswertung Die topologische Ähnlichkeitsauswertung wird lokal für jedes Knotenpaar entsprechend dem von Zager und

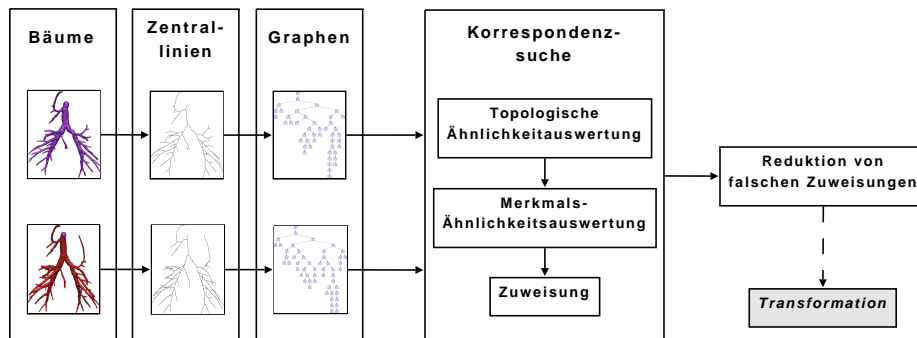


Abb. 1. Schematischer Ablauf für die Registrierung anatomischer Baumstrukturen am Beispiel eines Bronchialbaums.

Vergheese [5] vorgestellten gekoppelten Knoten-Kanten-Auswertungsverfahren berechnet. Bei gegebenen zwei Knoten $v_A \in V_A$ und $v_B \in V_B$, und den Funktionen $s(e)$, welche den Quellknoten einer Kante e liefert, und $t(e)$, welche den Zielknoten liefert, berechnet sich das Ähnlichkeitsmass (m) iterativ folgendermaßen:

$$m_{v_B v_A}^k \leftarrow \sum_{s(e_B)=v_B, s(e_A)=v_A} (m_{s(e_B)s(e_A)}^{k-1} + m_{t(e_B)t(e_A)}^{k-1}) + \sum_{t(e_B)=v_B, t(e_A)=v_A} (m_{s(e_B)s(e_A)}^{k-1} + m_{t(e_B)t(e_A)}^{k-1}) \quad (1)$$

bis $|m_{v_B v_A}^k - m_{v_B v_A}^{k-1}| \leq \varepsilon$ (ε : Konvergenzgrenze) mit $m_{v_B v_A}^0 = n_{v_B v_A}$ (vgl. Abschnitt 2.1).

Merkmals-Ähnlichkeitsauswertung Da viele Knoten topologisch identisch sind, ist allein die topologische Ähnlichkeit nicht ausreichend, um eine passende Zuweisung zu finden. Daher werden zusätzliche *Merkmale* wie der Gefäßdurchmesser für alle Knoten berechnet und deren Ähnlichkeiten ebenfalls bewertet. In dieser Arbeit wurden die folgenden sechs Merkmale verwendet, von denen keines Angaben über die absolute Position eines bestimmten Knotens benötigt, so dass das Zuweisungsverfahren lageunabhängig ist: Gefäßdurchmesser (mm), innerer Abstand zur Wurzel (mm), Knotenebene, maximaler und minimaler Winkel zwischen den Kanten, welche die Verbindung zu den Kinderknoten repräsentieren und der Winkel zwischen der virtuellen Verbindungslinie von der Wurzel zum berücksichtigten Knoten und der ersten von der Wurzel ausgehenden Kante. Die Ähnlichkeitsauswertung wird durch einen Gaußkern im Intervall $[0, 1]$ bewertet (1: identische Merkmalswerte). Sei p_v^P der Wert eines bestimmten Merkmals P eines Knotens v und $G(p_{v_B}^P, p_{v_A}^P)$ der Gaußkern, so berechnet sich das Merkmals-Ähnlichkeitsmaß folgendermaßen:

$$n_{v_B v_A} \leftarrow \sum_P G(p_{v_B}^P, p_{v_A}^P) \quad (2)$$

Zuweisung Zur Bestimmung von Korrespondenzen basierend auf den berechneten Ähnlichkeitswerten $w_{v_B v_A} = m_{v_B v_A} + n_{v_B v_A}$ verwenden wir den *Munkres* Algorithmus [6].

2.2 Reduktion von falschen Zuweisungen

Zur Reduktion von falschen Zuweisungen gehen wir davon aus, dass wenn ein bestimmter Knoten v_A einem Knoten v_B zugewiesen wurde, andere Knoten in der topologischen Nachbarschaft von v_A existieren müssen, welche entsprechenden Knoten in der Nachbarschaft von v_B zugewiesen wurden. Liegt die Anzahl solcher Knoten unter einem Schwellwert β , wird die Zuweisung gelöscht.

Tabelle 1. Ergebnisse der Registrierung von Bäumen mit K_1 und K_2 Knoten, von denen K_Z eine Korrespondenz der Referenzannotierung zugewiesen wurde und für K_U manuell keine Korrespondenz gefunden wurde. Richtig, falsch und nicht zugewiesene Knotenzahlen werden mit K_r , K_f und K_n notiert. Die Gesamtberechnungsdauer t umfasst die Arbeitsschritte der topologischen und Merkmals-Ähnlichkeitsauswertung, Zuweisung und Reduktion von falschen Zuordnungen.

	Datensatznummer										
	1	2	3	4	5	6	7	8	9	10	$\emptyset(\%)$
K_1	28	23	25	44	52	54	25	50	22	56	
K_2	28	20	16	29	45	35	24	44	17	43	
K_Z	24	19	15	27	42	35	19	36	15	41	91±7
K_r	20	11	9	13	30	22	9	19	9	22	60±11
K_f	0	2	0	2	1	0	0	2	0	2	3±4
K_n	4	6	6	12	11	13	10	15	6	17	37±10
K_U	4	1	1	2	3	0	5	8	2	2	9±7
K_r	4	1	0	2	2	-	5	7	2	2	84±33
K_f	0	0	1	0	1	-	0	1	0	0	16±33
t [ms]	20	$< 10^{-2}$	100	330	170	1810	20	80	20	210	

3 Ergebnisse

Zur Evaluation des vorgestellten Verfahrens wurden 15 CT-Datensätze von Schweinelungen untersucht, die jeweils aus einem Expirations- und einem Inspirationsvolumen bestehen. Die Bronchialbäume wurden segmentiert und in eine Graphdarstellung überführt. Anschließend wurden manuell korrespondierende Verzweigungen identifiziert, die später als Referenz dienten. Fünf der Datensätze wurden zur Optimierung der Parameter unseres Verfahrens (Parameter der Gaußkerne u. der Methode zur Reduktion von falschen Zuweisungen) genutzt, während die verbleibenden 10 Datensätze mittels der optimierten Parameter zur Evaluation der vorgestellten Methode verwendet wurden. Die Ergebnisse sind in Tabelle 1 aufgeführt. 60% aller zuweisbaren Knoten – das sind Knoten mit einer manuellen Referenz – wurden richtig zugewiesen (*zuweisbar, richtig*), während 3% falsch (*zuweisbar, falsch*) und 37% nicht zugewiesen wurden (*zuweisbar, unzugewiesen*). Unter den Knoten, welche nicht zuweisbar waren, wurden 16% falsch zugewiesen (*unzuweisbar, falsch*), während 84% korrekt als nicht zuweisbar identifiziert wurden (*unzuweisbar, richtig*). Zur Erzeugung und Merkmalsberechnung von zwei Bäumen benötigte das Verfahren weniger als 10^{-3} ms (2.4 GHz Intel CPU). Das gesamte Registrierungsverfahren einschließlich topologischer und Merkmals-Ähnlichkeitsauswertung, der Zuweisung und der Reduktion von falschen Zuordnungen beanspruchte durchschnittlich weniger als eine Sekunde.

4 Diskussion

In dieser Arbeit wurde ein schnelles und robustes Verfahren zur Korrespondenzsuche in baumförmigen Strukturen präsentiert, das sowohl topologische also auch anatomische Ähnlichkeit von Knoten nutzt, um eine global optimale Zuweisung zu berechnen. Vorteile des Verfahrens gegenüber anderen aus der Literatur bekannten Ansätzen sind die Möglichkeit einer *global* verteilten Zuordnung sowie die Unabhängigkeit gegenüber der relativen Lage der Bäume zueinander. Des Weiteren umfasst das Verfahren eine Methode zur Reduktion von falschen Zuweisungen.

Unsere Ergebnisse bezüglich Laufzeit und Anteil an zugewiesenen Knoten sind mit denen von Graham und Higgins [2, 3] vergleichbar. Allerdings unterscheiden die Autoren nicht zwischen falsch und richtig zugewiesenen Knoten. Die Laufzeit wurde hier nur für einen der vier untersuchten Datensätze angegeben (5 Sekunden für zwei Bäume mit 341 und 131 Knoten; PC 3.4 GHz Pentium). Im Vergleich zu der Arbeit von Metzen et al. [1] liefert unser Verfahren einen höheren Anteil an richtigen Zuweisungen (53% gegenüber 19%), während der Anteil an falsch zugewiesenen Knoten vergleichbar ist (4% gegenüber 2%). Laufzeitangaben zum Algorithmus wurden keine angegeben.

Zusammenfassend lässt sich feststellen, dass das vorgeschlagene Verfahren eine schnelle und robuste Möglichkeit darstellt, um anatomische Bäume zu registrieren und im Vergleich zum aktuellen Stand der Forschung einen hohen Anteil an richtig zugewiesenen Knoten sowie einen niedrigen Anteil an falsch zugewiesenen Knoten liefert.

Danksagung. Diese Arbeit entstand innerhalb des Stipendiums des CAPES/DAAD (Brasilien/Deutschland) Austauschprogramms von Thiago R. dos Santos.

Literaturverzeichnis

1. Metzen JH, Kröger T, Schenk A, et al. Matching of anatomical tree structures for registration of medical images. *Image Vis Comput.* 2009;27:923–33.
2. Graham MW, Higgins WE. Globally optimal model-based matching of anatomical trees. In: *Proc SPIE*; 2006. p. 373–87.
3. Graham MW, Higgins WE. Optimal graph-theoretic approach to 3D anatomical tree matching. In: *Proc ISBI*; 2006. p. 109–12.
4. Wolber P, Wegner I, Heimann T, et al. Tracking und Segmentierung baumförmiger, tubulärer Strukturen mit einem hybriden Verfahren. In: *Proc BVM*; 2008. p. 242–6.
5. Zager LA, Verghese GC. Graph similarity scoring and matching. *Appl Math Lett.* 2008;21:86–94.
6. Bourgeois F, Lassalle JC. An extension of the Munkres algorithm for the assignment problem to rectangular matrices. *Commun ACM.* 1971;14(12):802–4.

Mosaickingalgorithmus zur schnellen Panoramabilderstellung in der Fluoreszenzendoskopie

Alexander Behrens¹, Michael Bommers¹, Thomas Stehle¹, Sebastian Gross¹,
Steffen Leonhardt², Til Aach¹

¹Lehrstuhl für Bildverarbeitung, RWTH Aachen University

²Philips Lehrstuhl für Medizinische Informationstechnik, Helmholtz-Institut,
RWTH Aachen University

`alexander.behrens@lfb.rwth-aachen.de`

Kurzfassung. Bisherige Mosaickingalgorithmen zur Erstellung von Panoramabildern der Harnblase aus endoskopischen Fluoreszenzbildern erreichen aufgrund ihres hohen Rechenaufwands nur einen langsamen sukzessiven Panoramabildaufbau. Die Verwendung eines Panoramas als interagierende Orientierungs- und Navigationshilfe erfordert jedoch eine echtzeitnahe Bilderstellung während der Operation. Hierzu wurde ein PDD-Mosaickingalgorithmus entwickelt, der eine schnelle Erstellung und Darstellung von endoskopischen Panoramabildern durch parallelisierte und synchronisierte Multi-Thread Prozesse ermöglicht.

1 Einleitung

Die Sichtung und Therapie von multi-fokalem Harnblasenkrebs erfolgt durch eine Zystoskopie, in der ein starres Endoskop durch die Harnröhre in die Harnblase des Patienten eingeführt wird. Bei Verwendung eines Videoendoskopiesystems wird hierbei Licht durch ein Lichtleiterkabel in die Blase eingestrahlt und ein Videobild durch einen am Okular arretierten Kamerakopf aufgenommen. Während unter Weißlicht der Kontrast zwischen kleinen flachen Tumoren und dem umliegenden gesunden Gewebe gering ist, liefert die photodynamische Diagnose (PDD) eine deutliche Kontrastverstärkung. Hierbei wird unter Verwendung einer Markersubstanz und einer schmalbandigen blauen Lichteinstrahlung (380–450nm) Tumorgewebe zum rötlichen Fluoreszieren angeregt. Zur Erzeugung von gut belichteten und kontrastreichen PDD-Bildern ist aufgrund der schmalbandigen Beleuchtung das Endoskop stets nah ($\leq 1\text{cm}$) an der Blasenwand entlangzuführen. Dies führt zu Ansichten mit einem sehr geringen „Field Of View“ (FOV), welches die Orientierung und Navigation innerhalb der Harnblase erschwert.

Panoramabilder iterativer Mosaickingalgorithmen [1, 2, 3] liefern dagegen ein erweitertes Sichtfeld. Obwohl offline berechnete Panoramabilder und deren Darstellung zur Befundsprotokollierung und Operationsplanung [1] verwendet werden können, ist für eine interagierende Navigationshilfe eine echtzeitnahe Verarbeitung sicherzustellen. Hierzu wird im Folgenden ein neuer PDD-Mosaickingalgorithmus vorgestellt, der durch ein Multi-Thread Softwaredesign innerhalb des

RealTimeFrame [4] eine schnelle Panoramabilderstellung ermöglicht. Mit Hilfe eines herkömmlichen PC-Systems wird hierbei bei voller Bildauflösung (768x576) das sukzessiv wachsende lokale Blasenpanorama live berechnet und direkt während der Blasen Spiegelung dargestellt.

2 Methoden

2.1 RealTimeFrame

Für die Entwicklung des Mosaickingalgorithmus wurde die Demonstratorplattform *RealTimeFrame* [4] verwendet. Basierend auf einem Model-View-Controller Konzept stellt das Framework Laufzeitinformationen und ein Hauptfenster zur Livebildanzeige und GUI-Elemente für Benutzereingaben bereit. Der Controller verwaltet Threads zur Bildakquisition und zur sequentiellen Abarbeitung einer konfigurierbaren Prozesskette. Algorithmen kompiliert als Dynamic Link Libraries können in die Prozesskette eingefügt und direkt ausgeführt werden. Dies ermöglicht ein schnelles Umsetzen von Echtzeit-Bildverarbeitungsalgorithmen [5], wie z.B. Kontrastverstärkung, zeitliche Filterung und Gefäßverstärkung.

2.2 Mosaickingalgorithmus

Die Erstellung eines lokalen PDD-Blasenpanoramas erfolgt durch eine Weiterentwicklung des iterativen Mosaickingalgorithmus aus [1, 2]. Hierbei wird zunächst der elliptische Vordergrundbereich in den Endoskopbildern extrahiert. Aufgrund der deutlich abgebildeten Gefäßstruktur der Blasenwand, werden danach einzelne charakteristische Merkmalspunkte durch „Speeded Up Robust Features“ (SURF) [6] extrahiert, welche eine sehr große Wiederholbarkeit in den Bildsequenzen aufweisen. Das Matchen von Punktkorrespondenzen zweier sequentieller Bilder erfolgt mittels der Featuredeskriptoren \mathbf{d} . Als Distanzfunktion wird der minimale euklidische Abstand

$$\min_j \|\mathbf{d}_i - \mathbf{d}_j\|_2 = \min_j \sqrt{\sum_{k=1}^n (d_{i,k} - d_{j,k})^2}, \quad (1)$$

zwischen den n -dimensionalen Deskriptoren $\mathbf{d}_i, \mathbf{d}_j$ herangezogen. Die anschließende Berechnung der Bild-zu-Bild-Homographie

$$\mathbf{H} = \begin{bmatrix} \mathbf{A} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad \text{mit} \quad \mathbf{A} = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \quad (2)$$

beschreibt eine affine Transformation mit vier Freiheitsgraden (Rotationswinkel α , Skalierung s_x, s_y , Scherung a) und eine Translation \mathbf{t} . Zur robusten Homographieschätzung wird ein adaptiver RANSAC (RANDOM SAMPLE CONSENSUS) Algorithmus [7] verwendet, der falsche Punktkorrespondenzen verwirft. Nach dem Ausführen der Bildtransformation wird abschließend der entstandene Überlappungsbereich mittels einer linearen Interpolation überblendet und ein Kompositionsbild erstellt. Iterativ werden so nachfolgende Videobilder sukzessiv dem Panoramabild hinzugefügt.

2.3 Softwaredesign

Zur schnellen Berechnung und Darstellung des entstehenden Panoramabildes während der Zystoskopie verfügt der Mosaickingalgorithmus, eingebettet in dem *RealTimeFrame*-Framework, eine Multi-Thread Softwarearchitektur und ein adaptives Bufferkonzept. Diese und eine Ablaufbeschreibung der Threads sind in Abb. 1 skizziert. Im ersten Prozessschritt wird das aktuelle Endoskopbild F_n dem *CentralBuffer* übergeben. Dieser verwaltet dafür K *PictureBuffer* bestehend aus jeweils N freien Speicherslots, welche sequentiell gefüllt werden. Vollständig gefüllte Buffer werden von der *EmptyBufferList* in die *ActiveBufferList* übertragen und zur Verarbeitung freigegeben. Da während der Blasensichtung die Freihandbewegung des Endoskops annähernd gleichmäßig und überwiegend langsam ist, weisen zwei direkt aufeinanderfolgende Bilder meist einen unnötig großen Überlappungsbereich auf. Dagegen führen größere und adaptiv gesteuerte Schrittweiten zwischen zwei zu registrierenden Bildern zu einem geringeren Rechenaufwand ohne das erweiterte FOV des Panoramabildes zu beeinflussen. Andererseits ergibt sich eine erhöhte Latenz durch das Buffern von Zwischenbildern, wobei die Anzahl der Speicherslots N gleich der Schrittweite ΔF ist.

Zunächst werden nur das erste und das N -te Endoskopbild zur Verarbeitung markiert und in die *NextElementList* eingereiht. Sobald diese Liste gefüllt wird, aktiviert sich der *DescriberThread*, extrahiert SURF Features und übergibt das Bild samt Deskriptoren d an den *ProcessedPictureBuffer*. Nun berechnet der *MatcherThread* Punktkorrespondenzen und schätzt daraus eine Bild-zu-Bild-Homographie H nach Gl. 2. Diese wird abgespeichert und der zu den Bildern F_1, F_N zugehörige *PictureBuffer* zum erneuten Füllen in die *EmptyBufferList* verschoben, da die Zwischenbilder F_2, \dots, F_{N-1} nicht weiter berücksich-

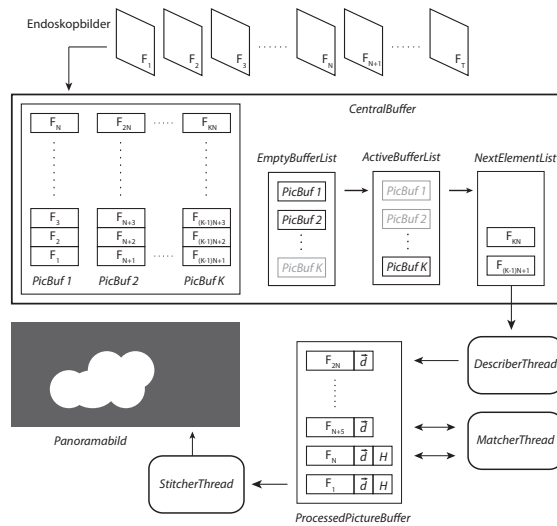


Abb. 1. Softwarearchitektur, Bufferdesign und Thread-Ablaufbeschreibung.

tigt werden müssen. Werden weniger als drei Punktkorrespondenzen bestimmt, so schlägt die Berechnung der Homographie fehl und der *MatcherThread* fordert ein neues Zwischenbild F_n mit $F_1 < F_n < F_N$ aus dem *PictureBuffer* an. Sobald eine Homographie vorliegt, fügt der *StitcherThread* das zugehörige Bild an das sukzessiv wachsende Panoramabild an und interpoliert die überlappenden Bildbereiche.

3 Ergebnisse

Das erzeugte lokale Panoramabild einer realen Zystoskopievideosequenz (609 Frames) ist in Abb. 2(a) dargestellt. Dieses bietet ein vergrößertes FOV der Blasenwand und zeigt zwei papillare Tumore an der linken oberen Blasenwand vom linken Ureterostium. Aufgrund der parallelisierten Prozesse (Abb. 1), bestimmt die Threadssynchronisation die Gesamtzeit zur Berechnung und Aktualisierung des Panoramabilds. Abb. 2(b) zeigt die durchschnittliche Prozesszeit der drei Threads über der Featurepunktanzahl P , gemessen auf einem Standard PC-System mit 2.2 GHz Intel Xeon Dual Prozessoren. Die Prozesszeit des Describers ist hierbei linear abhängig von P , da der Rechenaufwand eines Featuredeskriptors d konstant ist. Der Offsetwert ist durch die vom Bildinhalt unabhängigen Vorverarbeitungsschritte wie Multiskalenrepräsentation, Filteroperationen, usw. bestimmt. Die Prozesszeit des Matchers dagegen erhöht sich quadratisch mit $\frac{P(P-1)}{2}$. Der Stitchingprozess ist zeitlich konstant und unabhängig von P , da ausschließlich die Bildtransformation H und das Blending ausgeführt wird. Der Verlauf der Gesamtzeit zwischen Bildeingang und Panoramabildanzeige ist in Abb. 2(b) als Summation der Einzelprozesszeiten und durch eine sequentielle

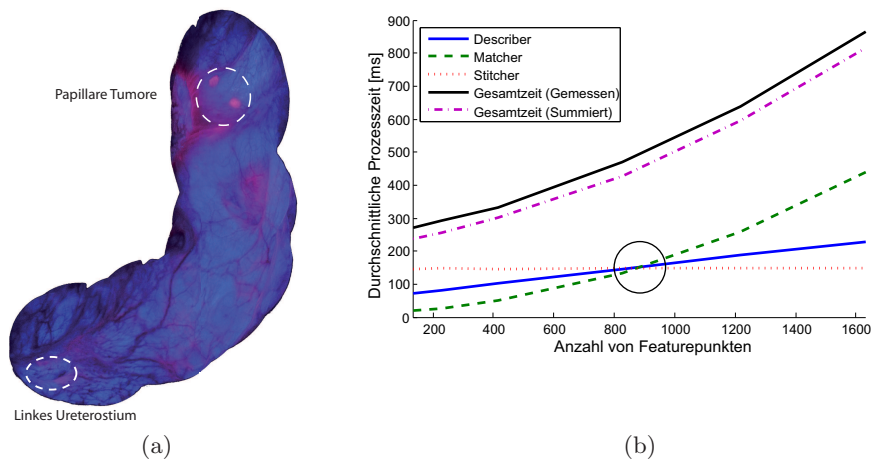


Abb. 2. Blasenpanorama bestehend aus 32 Einzelbildern (a) und durchschnittliche Prozesszeit der Threads über der Anzahl von Featurepunkten (b).

Messung dargestellt. Die konstante Abweichung beschreibt die Rechenzeit des Speicher-Managements und allgemeinen Systemoperationen.

4 Diskussion

Der optimale Kompromiss zwischen Rechenzeit und großer Featureanzahl P ist bestimmt durch die Thread-Synchronisation. Ein Optimum wird bei gleicher Prozessdauer der parallelisierten Threads erreicht. Dieser Arbeitspunkt ist in Abb. 2(b) durch den Schnittpunkt der Zeitverläufe bei $P_A = 855$ bestimmt. Eine Reduzierung der Gesamt-Rechenzeit ist somit direkt abhängig von den einzelnen Prozesszeiten der drei Threads. Nach Optimierung der Bildschrittweite ΔF in dem Arbeitspunkt P_A , beträgt die Gesamtverarbeitungszeit für ein Bild mit voller Auflösung (768×576) und Panoramadarstellung ohne jegliche Abwärtsabtastung [1, 2] im Durchschnitt 257.9 ms. Die Bildwiederholrate ergibt sich damit zu 3.8 Bilder/Sekunde. Der Multi-Thread-basierte Mosaickingalgorithmus ermöglicht somit eine schnelle sukzessive Erstellung von PDD-Bildkompositionen. Bei einer typischen Zystoskopbewegung, liefert der Algorithmus ein robustes Bildstitching, da der Überlappungsbereich zweier Bildpaare ausreichend groß ist. Aufgrund der erreichten Bildwiederholrate kann das Panoramabild dem Chirurg nun während der Blasenspiegelung auf einem zweiten Monitor als zusätzliche Orientierungshilfe dargestellt und erste klinische Studien erhoben werden. Weitere Prozessanalysen der parallelisierten Frameworks können unabhängig voneinander erfolgen und neue Feature-Detektoren, Matching- oder Blending-Methoden implementiert und optimiert werden. Darüber hinaus können zusätzliche Prozesse wie iterative Registrierungsverfahren (z.B. Bundleadjustment [3]) in Form von neuen Threads dem System schnell und einfach hinzugefügt werden.

Literaturverzeichnis

1. Behrens A, Stehle T, Gross S, et al. Local and global panoramic imaging for fluorescence bladder endoscopy. In: Proceedings of EMBC; 2009. p. 6990–3.
2. Behrens A. Creating panoramic images for bladder fluorescence endoscopy. Acta Polytechnica Journal of Advanced Eng. 2008;48(3):50–4.
3. Miranda-Luna R, Daul C, Blondel W, et al. Mosaicing of bladder endoscopic image sequences: distortion calibration and registration algorithm. IEEE Trans Biomed Eng. 2008;55(2):541–553.
4. Gross S, Stehle T. RealTimeFrame - A real time processing framework for medical video sequences. Acta Polytechnica Journal of Advanced Eng. 2008;48(3):15–9.
5. Gross S, Behrens A, Stehle T. Rapid development of video processing algorithms with RealTimeFrame. In: Biomedica. Liege; 2009. p. 217–20.
6. Bay H, Ess A, Tuytelaars T, et al. SURF: Speeded Up Robust Features. Comput Vis Image Underst. 2008;110(3):346–59.
7. Fischler M, Bolles R. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Commun ACM. 1981;24(6):381–95.

Günstige Kamerapfade für medizinische Animationen

Konrad Mühler, Bernhard Preim

Institut für Simulation und Graphik, Otto-von-Guericke-Universität Magdeburg
muehler@isg.cs.uni-magdeburg.de

Kurzfassung. Bei der Planung von Operationen spielen interaktive Animationen eine wichtige Rolle. Zur ihrer Unterstützung wurde ein Verfahren zur verbesserten Planung von Kamerapfaden entwickelt, wobei besonderen Wert auf die Vermeidung ungewohnter Sichten und die Erhöhung der Informationen während der Kamerafahrt gelegt wurde. Außerdem konnte durch ein Verfahren zum Clustering von Kamerapositionen eine Reduzierung der anzufahrenden Positionen erreicht werden.

1 Einleitung

Animationen finden immer mehr Anwendung bei der chirurgischen Operationsplanung, der Dokumentation und Präsentation. Dabei werden zumeist 3D-Volumendarstellungen kritischer Strukturen genutzt, um räumliche Ausdehnungen und Zusammenhänge zwischen den Strukturen besser beurteilen oder erkennen zu können. Animationen können den Chirurgen bei der Exploration solcher Darstellungen direkt unterstützen, indem bspw. Übergänge zwischen zwei Sichten oder Darstellungsformen animiert werden. Dies erhöht die Übersicht und fördert die Orientierung des Chirurgen. Auch Animationen in Form automatisch generierter Videos sind hilfreich, wenn es darum geht, kritische Fälle in einer interdisziplinären Besprechung kompakt vorzustellen und interdisziplinär zu diskutieren. Ein drittes Anwendungsgebiet für medizinische Animationen ist der Bereich der Ausbildung. Hier werden bspw. mit Hilfe von Animationen einzelne Fälle aufbereitet präsentiert, unterstützt durch verschiedene Techniken wie textuelle Annotationen [1]. Bei der Erstellung von Animationen für den klinischen Alltag sind drei Aspekte von zentraler Bedeutung: 1) Die Animationen müssen möglichst automatisch erstellt werden, 2) Es müssen gute Sichtpunkte auf kritische Strukturen oder die gesamte Darstellung automatisch ermittelt werden und 3) die einzelnen Sichten müssen angemessen durch Kamerapfade verbunden werden. Die Punkte 1) und 2) wurden bisher schon gründlich untersucht. Wir widmen uns daher dem noch ungelösten Problem der Ermittlung angemessener Kamerapfade zur Verbindung guter Sichtpunkte.

1.1 Stand der Forschung

Es sind verschiedene Vorarbeiten zur Kamerapfadplanung in 3D-Szenen bekannt. Sokolov et al. [2] berechnen einen Pfad um eine Szene, der möglichst viele gute

Sichtpunkte enthält. Iserhardt-Bauer et al. [3] zeigen ein geclipptes 3D-Volume-rendering aus festen Sichten, die sie über standardisierte Pfade miteinander zu einem Video verbinden. Sokolov und Plemenos [4] zeigen während einer Kamerafahrt durch eine Szene ein Objekt solange, bis alle wesentlichen Informationen gezeigt wurden und schwenken dann zum nächsten Objekt. In [5] wurde ein Verfahren zur automatischen Kamerapositionierung für einzelne Strukturen in einer Darstellung vorgestellt. Dabei fließen unterschiedliche Parameter wie die sichtbare Fläche und die Wichtigkeit verdeckender Strukturen in die Berechnung einer optimalen Kameraposition ein. Jeder Parameter wird durch eine *parameter map* verkörpert, die die Werte eines Parameters für alle möglichen Kamerapositionen enthält. Zur Ermittlung der besten Kameraposition werden die einzelnen *parameter maps* gewichtet summiert. Das Maximum des Summenfeldes ist die neue Kameraposition.

2 Methoden

Die animierte Bewegung der Kamera in einer 3D-Darstellung kann in verschiedenen Situationen notwendig sein. So ist in einer automatisch wiedergegebenen Animation die Bewegung der Kamera offensichtlich. Aber auch während der Interaktion kann eine automatische Veränderung der Kameraposition erforderlich sein, wenn bspw. der Nutzer eine Struktur selektiert, woraufhin die Kamera automatisch zu einem guten Sichtpunkt auf diese Struktur bewegt wird. Dies spart dem Nutzer Zeit, da er nicht aufwendig selbst in der Darstellung navigieren muss. Ein Kamerapfad gilt aus unserer Sicht als angemessen, wenn er folgende Kriterien erfüllt: a) Der Betrachter muss zu jeder Zeit wissen, wo sich die Kamera gerade befindet, b) der Pfad sollte möglichst kurz sein, c) die Bewegung der Kamera sollte flüssig und angenehm sein, d) während der Kamerafahrt sollten keine ungewohnten Sichten eingenommen werden, e) es sollte stets ein Maximum an Informationen präsentiert werden. Um diesen Kriterien gerecht zu werden, haben wir ein existierendes Animations-Framework [6] um dezidierte Techniken zur automatischen Kamerapfadplanung erweitert.

2.1 Erhalt der Orientierung

Werden zwei Sichtpunkte animiert verbunden, so sind auch die dazwischen gezeigten Sichten von essenzieller Bedeutung. Wird in beiden Sichten sehr nah an die Strukturen herangezoomt, so muss während der Kamerafahrt herausgezoomt werden, um die Übersicht zu behalten. Ein Herauszoomen verlängert eine Kamerafahrt jedoch unnötig, wenn die beiden Sichtpunkte sehr nah beieinander liegen. Die Stärke des Herauszoomens ist daher abhängig von der zurückzulegenden Strecke. Sie hat ein Maximum bei einer Höhe, von der aus die gesamte Szene zu erkennen ist. Bei längeren Kamerafahrten wird zu Beginn auf die Maximalhöhe herausgezoomt, um anschließend für den weiteren Teil der Kamerafahrt die gesamte Szene als Orientierung zu sehen. Abschließend wird auf die Zielstruktur in die Szene gezoomt.

2.2 Orientierung an einer Vorzugsblickrichtung und den Sichtbarkeiten von Strukturen

Untersuchungen haben gezeigt, dass Ärzte bei der Exploration von 3D-Szenen bestimmte Blickrichtungen bevorzugen [5]. Diese gilt es bei Kamerafahrten zu berücksichtigen. So wird bspw. bei HNO-Operationsplanungen vermieden, dass die Kamera nah zu den Polen der szenenumgebenden Kugel bewegt wird, und so für den HNO-Chirurgen ungewohnte und somit störende Sichten entstehen. Das vorliegende Verfahren berechnet eine Kamerafahrt unter Berücksichtigung eines Parameterfeldes, welches zu jeder Kameraposition auf der szenenumgebenden Kugel einen Gütwert im Bezug auf die Vorzugsblickrichtung enthält. Dies können hohe Werte entlang des Äquators der Kugel und niedrige Werte an den Polen sein (Abb. 1(a)). Vom Startpunkt aus wird in Zielrichtung in einem Öffnungswinkel von 45° nach der jeweils nächsten Position im Bezug auf den Wert im Parameterfeld der Vorzugsblickrichtung gesucht, bis die Zielposition erreicht ist (Abb. 1(b)). Eine Glättung des Pfades vermeidet ruckartige Richtungsänderungen und erzeugt eine flüssige Kamerabewegung.

Befinden sich am Beginn und Ende einer Kamerafahrt Strukturen, deren Sichtbarkeit für alle Kamerapositionen bekannt ist, so fließt auch diese in die Pfadberechnung mit ein. Es kann so zusätzlich garantiert werden, dass die Kamera während des Fluges noch eine Zeit lang die bisherige Struktur von Interesse und später die Zielstruktur sieht und diese nicht den gesamten Flug über verdeckt sind. Die möglichst frühe und permanente Sichtbarkeit der Zielstruktur ermöglicht dem Betrachter schon früh eine Orientierung im Bezug auf die Zielstruktur. Dazu werden die Parameterfelder der Sichtbarkeitswerte mit dem Parameterfeld der Vorzugsblickrichtung gewichtet summiert. Ein Pfad, der nach dem oben stehenden Prinzip generiert wird, garantiert so zusätzlich eine mög-

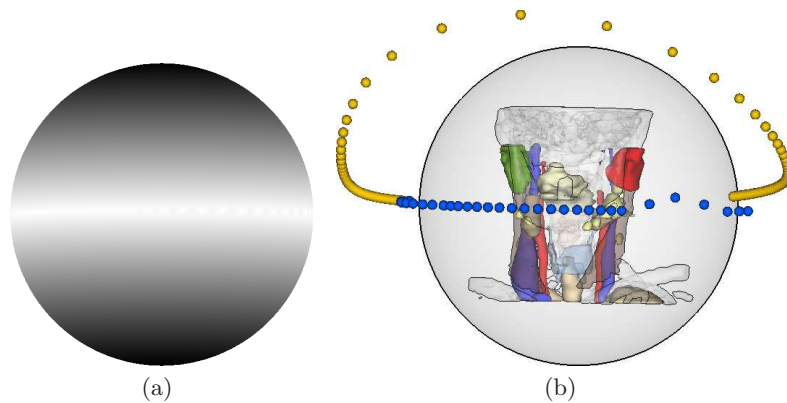


Abb. 1. (a): Parameterfeld einer Vorzugsblickrichtung, bei der Kamerapositionen am Äquator als beliebte Sichtrichtungen definiert sind. (b): Vergleich eines optimierten Kamerapfades (blau) mit dem kürzesten Weg (gelb). Der kürzere Weg führt nah an ungewünschten Sichten entlang der Pole, während der optimierte Pfad sich an der Vorzugsblickrichtung sowie den Sichtbarkeiten der Start- und Zielstruktur orientiert.

lichst gute Sichtbarkeit der Strukturen von Interesse. Eine höhere Wichtung der Vorzugsblickrichtung sorgt dafür, dass dennoch keine ungewöhnlichen Sichten eingenommen werden.

2.3 Clustering von Kamerapositionen

Sollen mit der Kamera viele Strukturen nacheinander aus einer guten Sicht gezeigt werden, kann der gesamte Pfad schnell zu lang werden. Daher ist es angebracht, Sichtpunkte in der Art zusammenzufassen, dass möglichst mehrere Strukturen von einem Sichtpunkt aus gleichzeitig betrachtet werden können. Im Rahmen des vorgestellten Frameworks kommt dabei folgender Algorithmus zur Anwendung (Alg. 2):

- Grundlage des Verfahrens sind die *parameter maps* der Sichtbarkeit jeder Struktur, wie sie in [5] eingeführt wurden und die für jede Kameraposition auf einer umgebenden Kugel die Sichtbarkeit der Struktur gespeichert haben.
- Diese *parameter maps* werden an einem Schwellenwert von 70% Sichtbarkeit binarisiert, um nur noch solche Positionen mit einer Mindestsichtbarkeit zu erhalten (sichtbar=1, nicht sichtbar=0).
- Anschließend werden die Strukturen in einer Gruppe zusammengefasst, bei denen nach der Multiplikation ihrer binarisierten *parameter maps* noch mindestens eine Position mit 1 bewertet wird. Dies bedeutet, dass mindestens eine Position existiert, von der aus alle Strukturen einer Gruppe aus zu mindestens 70% sichtbar sind (Abb. 2).

3 Ergebnisse und Diskussion

Zur Unterstützung der automatischen Generierung von Animationen in medizinischen Visualisierungen wurde ein Verfahren zur verbesserten Planung von effizienten Kamerapfaden vorgestellt. Dabei wurde besonderen Wert auf die Vermeidung ungewohnter Sichten und die Erhöhung der Informationsdichte während des Fluges gelegt. Zusätzlich wurde durch ein Verfahren zum Clustering von Kamerapositionen erreicht, dass bei vielen zu inspizierenden Strukturen der Pfad

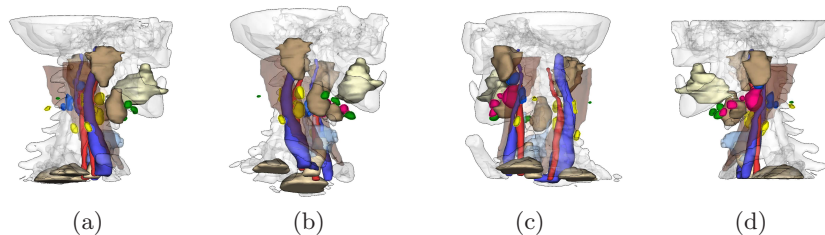


Abb. 2. Vier Sichten, die alle 18 Lymphknoten einer Visualisierung zur Planung von Lymphknotenausräumungen zeigen. Die Reduzierung von 18 auf 4 Sichten verkürzt Animationen erheblich und reduziert die Anzahl von Screenshots zur Dokumentation.

Algorithmus 2 Algorithmus zum Clustering von Kamerapositionen.

Input: Alle Parameterfelder PM**Output:** n Listen von Strukturen S die jeweils zusammen sichtbar sind und dazugehörige n Parameterfelder CLUSTER

tresh=0.7; {Schwellenwert zur Binarisierung}

n=0; {Initialisierung des Counters zum Zählen der Cluster}

for all $PM_i \in PM$ **do**PMBIN $_i \leftarrow bin(PM_i, \text{tresh})$ {Binarisierung der PM}

found = false;

for all CLUSTER $_j \in CLUSTER$ **do****if** $1 \in PMBIN_i * CLUSTER_j$ **then**CLUSTER $_j \leftarrow PMBIN_i * CLUSTER_j$;Füge i in die Liste S_j ein;

found = true;

end if**end for****if not found then**CLUSTER $_{n+1} \leftarrow PMBIN_i$;Füge i in die Liste S_{n+1} ein;**end if****end for**

erheblich gekürzt wird. Das Verfahren ist in mehreren Planungs- und Trainingssystemen integriert und erlaubt sowohl unterstützt dort neben der automatischen Videogenerierung auch das interaktive Pausieren einer Animation für mögliche individuelle Explorationen und eine anschließende Fortsetzung des Kamerafluges.

Danksagung. Dieses Projekt wurde durch das BMBF im Zusammenhang mit dem SOMIT-FUSION-Projekt gefördert (FKZ 01-BE03B). Wir danken MeVis Medical Solutions Bremen für die Bereitstellung von MEVISLAB.

Literaturverzeichnis

1. Mühler K, Preim B. Automatic textual Annotation for surgical planning. In: Proc VMV; 2009. p. 277–84.
2. Sokolov D, Plemenos D, Tamine K. Viewpoint quality and global scene exploration strategies. In: Proc GRAPP; 2006. p. 184–91.
3. Iserhardt-Bauer S, Rezk-Salama C, Ertl T, et al. Automated 3D video documentation for the analysis of medical data. In: Proc BVM; 2001. p. 409–13.
4. Sokolov D, Plemenos D. Virtual world explorations by using topological and semantic knowledge. Visual Computer. 2008;24(3):173–85.
5. Mühler K, Neugebauer M, Tietjen C, et al. Viewpoint selection for intervention planning. In: Proc EuroVis; 2007. p. 267–74.
6. Mühler K, Bade R, Preim B. Adaptive script based animations for intervention planning. In: Proc MICCAI; 2006. p. 478–85.

Intensitätsbasiertes Multiskalen-Blending zur Erstellung von Panoramabildern in der Fluoreszenzendoskopie

Alexander Behrens, Martin Guski, Thomas Stehle, Sebastian Gross, Til Aach

Lehrstuhl für Bildverarbeitung, RWTH Aachen University, 52056 Aachen, Germany
alexander.behrens@lfb.rwth-aachen.de

Kurzfassung. Bei der Erstellung von Bildkompositionen aus endoskopischen Fluoreszenzbildern kommt es bei herkömmlichen linearen Blendingverfahren aufgrund starker Bildintensitätsunterschiede zwischen den zu überblendenden Endoskopbildern zu ungewollten visuellen Artefakten. Das entwickelte nichtlineare intensitätsbasierte Pyramidenblending dagegen erhält sowohl diagnostisch relevante Fluoreszenzinformationen als auch feine Gefäßstrukturen. Dies ermöglicht die Erstellung von kontrast- und struktureicheren Panoramabildern der Harnblase.

1 Einleitung

Die medizinische Diagnose und Therapie von Harnblasenkrebs erfolgt durch eine Zystoskopie, in der ein starres Endoskop durch die Harnröhre in die Harnblase des Patienten eingeführt wird. Mittels eines Videoendoskopiesystems wird hierbei die Harnblase gesichtet. Zur Kontrastverstärkung zwischen gesunden und tumorösen Gewebe kommt die photodynamische Diagnose (PDD) zur Anwendung, bei der eine Markersubstanz angeregt durch schmalbandiges eingestrahktes blaues Licht (380–450 nm) fluoresziert. Aufgrund der schmalbandigen Lichteinstrahlung ist für die Erzeugung von rauscharmen und kontrastreichen Bildern das Zystoskop stets dicht (≤ 1 cm) an der Blasenwand entlangzuführen. Dies führt zu Bildern mit einem kleinen „Field of View“ (FOV), welches die Navigation und Orientierung innerhalb der Blase erschwert.

Mosaikalgorithmien [1, 2, 3] liefern dagegen durch iteratives Aneinandersetzen von sequentiellen Endoskopbildern lokale Blasenpanoramen mit größeren FOVs. Klinisch können diese zur Orientierungshilfe, zur Befundsdokumentation und zum gezielten Wiederauffinden von Tumoren eingesetzt werden. Hierfür ist jedoch sicherzustellen, dass die originale Bildinformation möglichst unverfälscht und ohne störende Artefakte im resultierenden Panoramabild wiedergegeben wird. Neben einer ungenauen Bildregistrierung kann ein unangepasstes Blendingverfahren starke visuelle Artefakte und Strukturverluste in den Überlappungsbereichen zweier überblendeter Bilder verursachen.

Obwohl lineare Alpha-Blendingmethoden [4] unter gleichbleibenden Beleuchtungsbedingungen gute Ergebnisse liefern [1, 2, 5], verschwinden bei Mehrfachüberblendungen von ungleich beleuchteten und aufgelösten Bildregionen dia-

gnostisch relevante Bildinformationen wie Fluoreszenzstärke und feine Vaskularisierungsstruktur. Um diese Informationen in den resultierenden Panoramabildern zu erhalten, wird im Folgenden ein neues intensitätsbasiertes Multiskalen-Blending für die Fluoreszenzendoskopie vorgestellt.

2 Methoden

Bedingt durch die Bauweise des optischen Systems des Zystoskops ist stets eine deutliche Vignettierung vorhanden, welche die Abschattung von der Bildmitte zum Bildrand beschreibt. Beim Überblenden des Überlappungsbereichs $A \cap B$ zweier Endoskopbilder A und B (Abb. 1(a)) ist daher die geringere Intensität der Randpixel zu berücksichtigen, um Artefakte wie in Abb. 1(b) gezeigt zu vermeiden. Unter Annahme von zwei annähernd gleich stark belichteten PDD-Bildern und eines radialsymmetrischen Vignettierungsverlaufs liefert ein Alpha-Blending mit einer linearen Gewichtungsfunktion, basierend auf der relativen Entfernung zwischen Pixelposition und Bildrand [5], bereits einen sanften Überblendungsverlauf ohne visuell störende Artefakte [1, 2] (Abb. 1(c)).

Eine annähernd gleichbleibende globale Beleuchtungsstärke der Blasenwand ist während eines Blasenscans durch die Freihandbewegung des Endoskops nicht sichergestellt. Aufgrund der schmalbandigen Lichteinstrahlung führen variierende Abstände zwischen Endoskopspitze und Blasenwand schnell zu einer deutlichen Intensitätsabnahme der Bildpixel (Abb. 2). Hohe Bildintensitäten dagegen deuten auf hochwertige und diagnostisch relevante Bildinformationen hin, da Gefäßstrukturen durch den geringeren Objektstand höher aufgelöst werden und stärkere Fluoreszenzen auf Tumore hinweisen.

Im Falle der Überblendung von unterschiedlich stark belichteten und gleichzeitig verschieden hoch aufgelösten PDD-Bildern (Abb. 2) mittels eines Alpha-Blendings, kommt es durch die lineare Pixelinterpolation zum Verlust von Strukturinformation. Darüber hinaus können bei Mehrfachüberblendungen selbst einst starke Fluoreszenzen von Tumoren vollständig verloren gehen.

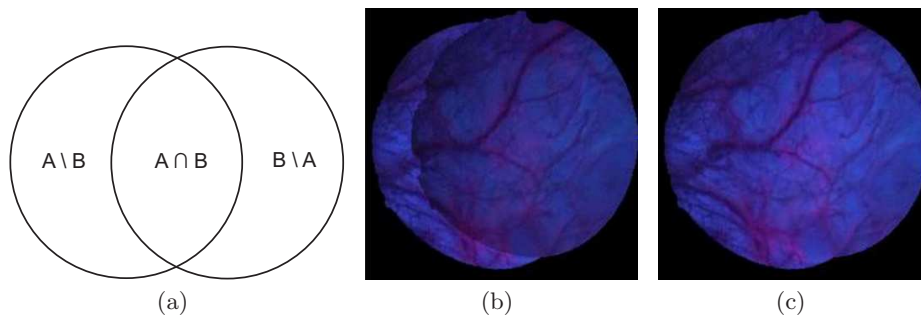


Abb. 1. (a) Überlappungsbereich zweier Endoskopbilder. (b) Bildregistrierung ohne Überblendung. (c) Bildregistrierung mit linear gewichtetem Alpha-Blending [1, 2, 5].

Ein ortsfrequenzabhängiges Blending [6] ermöglicht es dagegen unterschiedliche Helligkeiten anzugleichen und gleichzeitig feine Bildstrukturen zu erhalten. Basierend auf einer Multiskalenrepräsentation werden dabei tieffrequente Signale über weite Bildregionen überblendet, während hohe Ortsfrequenzen nur lokal berücksichtigt werden. Dieser Ansatz wird im Folgenden für die Erhaltung von feiner Vaskularisierungsstruktur in PDD-Panoramabildern übernommen. Dazu werden die Bildinhalte der beiden zu kombinierenden Endoskopbilder zunächst durch Laplacepyramiden [7] repräsentiert. Die Erstellung der Laplacepyramide erfolgt dabei rekursiv durch Gaußpyramiden

$$G^{(q+1)} = \downarrow_2 \mathcal{B} G^{(q)} \quad \text{mit} \quad G^{(0)} = G \quad (1)$$

wobei sich das Originalbild G auf der 0-ten Pyramidenebene befindet. Jede q -te Ebene berechnet sich mittels binomialer Tiefpassfilterung \mathcal{B} und anschließender Abwärtsabtastung \downarrow_2 um den Faktor 2. Die Laplacepyramide mit gewünschter Bandpasszerlegung ergibt sich damit zu

$$L^{(p)} = G^{(p)} - \mathcal{B} \uparrow_2 G^{(p+1)} \quad \text{mit} \quad L^{(P)} = G^{(P)} \quad (2)$$

in der das letzte Bild der Pyramide $L^{(P)}$ das tiefpassgefilterte Bild $G^{(P)}$ repräsentiert und \uparrow_2 einer Aufwärtsabtastung um den Faktor 2 entspricht.

Nach Erstellung der Laplacepyramiden L_A und L_B der Eingangsbilder A und B , wird eine Gaußpyramide der Blendingmaske des Überlappungsbereichs $A \cap B$ erstellt. Anstatt der Verwendung einer binären Maske [6] oder einer tiefpassgefilterten Randmaske [5] wird nun eine Gewichtungsfunktion verwendet, die sowohl die Vignettierung als auch die maximale Intensitätserhaltung in den PDD-Bildern berücksichtigt. Dazu werden zunächst Vignettierungsmasken M_A, M_B der Eingangsbilder erstellt, welche die Randabschattung des Endoskops durch eine streng monoton steigende radialsymmetrische Gewichtungsfunktion nachbilden. Die zum Blending verwendete Gewichtungsmaske

$$M = \begin{cases} M_A & \text{für } E_A \geq E_B \\ 1 - M_B & \text{für } E_A < E_B \end{cases} \quad \text{mit} \quad (3)$$

$$E_A = \sum_{A \cap B} (M_A \cdot I_A)^2 \quad \text{und} \quad E_B = \sum_{A \cap B} (M_B \cdot I_B)^2 \quad (4)$$

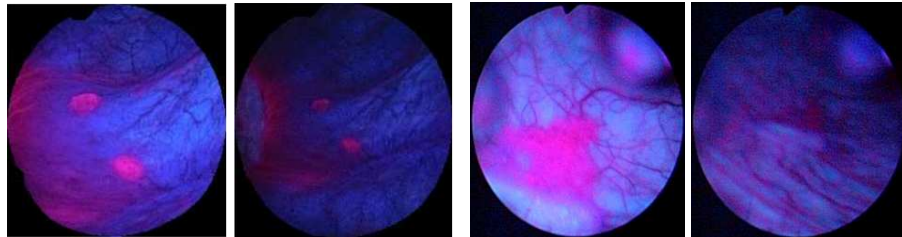


Abb. 2. Fluoreszierendes Tumorgewebe bei variierenden Abständen zwischen Endoskop und Blasenwand.

wird anschließend durch einen Vergleich der auftretenden Energien in beiden Grauwertbildern I_A, I_B aus den Vignettierungsmasken M_A, M_B gemäß Gl. 3 bestimmt. Zur abschließenden Bildrekonstruktion ist aufgrund der Pyramidenstruktur für die Maske M eine Gaußpyramide mit gleicher Stufenanzahl zu berechnen. Die rekursive Verknüpfung der mit M gewichteten Bandpassanteile von Bild A und B ergeben dann die resultierende Bildkomposition I_c^0 gemäß

$$I_c^{(p-1)} = M^{(p-1)} \cdot L_A^{(p-1)} + \left(1 - M^{(p-1)}\right) \cdot L_B^{(p-1)} + \mathcal{B} \uparrow_2 I_c^{(p)} \quad \text{mit (5)}$$

$$I_c^{(P)} = M^{(P)} \cdot L_A^{(P)} + \left(1 - M^{(P)}\right) \cdot L_B^{(P)}$$

Aufgrund der verwendeten 3-kanaligen Eingangsbilder ist die Rekonstruktion auf jedem der drei Farbkanäle gemäß Gl. (5) durchzuführen.

3 Ergebnisse

Zum Vergleich der vorgestellten Blendingverfahren wird eine reale, typische zystoskopische Freihandsequenz mit einer Auflösung von 384x288 verwendet, bei der während des Scans im Bereich zweier papillärer Tumore erst ein gewolltes Heranfahren und danach ein Wegbewegen des Endoskops von der Blasenwand vom Chirurg durchgeführt wurde (Abb. 2). Hierbei kommt es bei sequentiellen Mosaikalgorithmien [1, 2, 3] zum Mehrfachüberblenden von gleichen aber unterschiedlich stark beleuchteten Bildregionen. Aus Abb. 3(a) ist zu entnehmen, dass das linear gewichtete Alpha-Blending die in der Sequenz zunächst stark

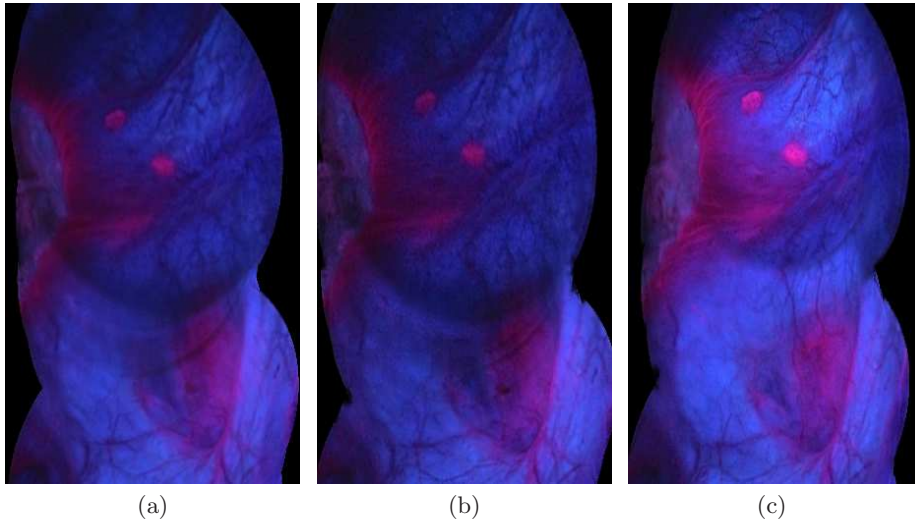


Abb. 3. (a) Lineares Alphablending, (b) lineares Pyramidenblending nach Burt [6], (c) nichtlineares Pyramidenblending.

fluoreszierenden Bildpixel durch Mehrfachüberblendungen sukzessive mit den intensitätsschwächeren Bildaufnahmen mittelt und feinere Vaskularisierungsstrukturen verloren gehen. Auch das lineare Multiskalen-Blending nach Burt [6] zeigt trotz Verwendung von Vignettierungsmasken M_A, M_B einen Verlust der Fluoreszenzinformation (Abb. 3(b)). Das hier entwickelte nichtlineare Pyramidenblending in Abb. 3(c) mit drei Pyramidenstufen erhält dagegen sowohl große Pixelintensitäten und damit stark fluoreszierende Bildregionen als auch feine Gewebestrukturen. Darüber hinaus werden weniger starke Kantenartefakte an den unvermeidlichen Helligkeitsverläufen zwischen dunklen und hellen Bildregionen erzeugt.

4 Diskussion

Lineare Alpha-Blendingmethoden erstellen bei annähernd konstant belichteten und ggf. normierten endoskopischen Videobildern sanfte Überblendungen ohne künstliche, visuell störende Artefakte. Bei starken Intensitätsschwanken zwischen den Bildern, wie sie in einer Freihandvideosequenz einer photodynamischen Diagnose vorkommen, erzeugen sie jedoch starke und unbefriedigende Blendingartefakte. Das entwickelte nichtlineare Pyramidenblending verwendet dagegen eine intensitätsabhängige Gewichtungsfunktion in unterschiedlichen Frequenzbändern. Hierbei wird erreicht, dass sowohl feine Strukturen als auch die Fluoreszenzintensität von tumorösen Gewebe für diagnostische Begutachtungen in einem Panoramabild der Harnblase erhalten bleiben. Dieses Verfahren ermöglicht die Erstellung von wesentlich struktureicheren Bildkompositionen für klinische Dokumentationsberichte und erlaubt es größere variierende Abstände zwischen dem Endoskop und der Blasenwand zu tolerieren, welche durch die Freihandbewegung des Chirurgen hervorgerufen werden können.

Literaturverzeichnis

1. Behrens A, Stehle T, Gross S, et al. Local and global panoramic imaging for fluorescence bladder endoscopy. In: Proc EMBC; 2009. p. 6990–3.
2. Behrens A. Creating panoramic images for bladder fluorescence endoscopy. Acta Polytech J Adv Eng. 2008;48(3):50–54.
3. Miranda-Luna R, Daul C, Blondel WCPM, et al. Mosaicing of bladder endoscopic image sequences: distortion calibration and registration algorithm. IEEE Trans Biomed Eng. 2008;55(2):541–553.
4. Porter T, Duff T. Compositing digital images. In: Proc SIGGRAPH; 1984. p. 253–259.
5. Wald D, Reeff M, Székely G, et al. Fließende Überblendung von Endoskopiebildern für die Erstellung eines Mosaiks. In: Proc BVM; 2005. p. 287 – 291.
6. Burt P, Adelson E. A multiresolution spline with application to image mosaics. ACM Trans Graph. 1983;2(4):217–36.
7. Jähne B. Digitale Bildverarbeitung. 6th ed. Springer; 2005.

Verfahren zur hochgenauen 3D-Rekonstruktion aus histologischen Schliffbildern

Waldemar Würfel¹, Andreas Hussong², Anna Herzog², Peter Erfurt¹,
Omid Majdani¹, Thomas S. Rau¹

¹Klinik und Poliklinik für HNO, Medizinische Hochschule Hannover

²Institut für Mechatronische Systeme, Leibniz Universität Hannover

rau.thomas@mh-hannover.de

Kurzfassung. Vorgestellt wird ein Verfahren zur präzisen Erstellung von 3D-Modellen anatomischer Strukturen aus äquidistanten, histologischen Schliffbildern. Dazu wird der Bildstapel zueinander registriert und ins DICOM-Format konvertiert. Der so entstandene Schliffbild-Datensatz bietet eine hochdetaillierte Auflösung und kann analog zu bzw. in Kombination mit konventionellen Bildgebungsverfahren verwendet werden. So wird die Segmentierung sehr feiner Weichgewebsstrukturen ermöglicht, die sich in herkömmlichen Modalitäten nicht oder nur schlecht abbilden. Erstmals wurde dabei die Rekonstruktionsgenauigkeit eines auf histologischen Schnitten basierenden Verfahrens anhand eines Referenzobjektes untersucht, und das Potential des vorgestellten Verfahrens validiert.

1 Einleitung

In den vergangenen Jahren treten zunehmend Anwendungsfälle auf, in denen die Forderung nach hoher Auflösung in Kombination mit der Notwendigkeit einer adäquaten Weichgewebsdifferenzierung mit aktuellen Bildgebungsverfahren wie Computertomographie (CT) oder Magnetresonanztomographie (MRT) nicht zufriedenstellend gelöst werden kann. Im Bereich der lateralen Schädelbasis ist z.B. im Kontext der Cochlea-Implantat-Versorgung die Darstellung und Bestimmung der Lage membranöser Strukturen der Hörschnecke (Cochlea) wie der Basilarmembran von großem Interesse. Verfügbare Mikro-CT ermöglicht nur bei Hochkontraststrukturen wie Knochen ausreichend detaillierte Aufnahmen; Weichgewebsstrukturen wie die Basilarmembran müssen erst aufwendig durch Einsatz von Kontrastmitteln der Bildgebung zugänglich gemacht werden [1].

Daher sind für die Erstellung detaillierter Anatomiemodelle mit submillimetrischer Genauigkeit schichtweise Präparationsverfahren das aktuelle Mittel der Wahl. Im Bereich der Modellierung des Mittelohres wurden dazu bereits Verfahren ausführlich beschrieben [2]. Ebenso verwendeten Sørensen et al. ein schneidendes Verfahren zur Gewinnung der Daten für das „Visible Ear Project“ [3, 4]. Besonders nennenswert in diesem Kontext ist zudem die Erstellung der „mouse cochlea database“ [5], bei dem durch ein optisches Verfahren (orthogonal-plane fluorescence optical sectioning microscopy, OPFOS) gewonnene Schichtbilder zu

einem 3D-Modell zusammengefügt wurden. Allen Verfahren ist jedoch gemein, dass die Genauigkeit der Rekonstruktion nicht untersucht wurde und zudem häufig komplexe chemische Vorbereitungsschritte einschließlich Dekalzifizierung notwendig sind, deren Einfluss auf die geometrische Gestalt der anatomischen Strukturen unbekannt ist. Bei dem hier vorgestellten, leicht zu reproduzierenden Verfahren kann durch die Verwendung einer Schlifftechnik auf den Schritt der Dekalzifizierung verzichtet werden. Zudem wurde die Methodik hinsichtlich der Rekonstruktionsgenauigkeit untersucht, für welche die Realisierung eines präzise bestimmbar Schichtabtrages notwendig ist. Hintergrund dieser Entwicklung ist das Ziel, aus diesen Schlifffbildern einen Anatomieatlas einschließlich der relevanter Weichgewebsanatomie zu erstellen, durch welchen die patientenindividuellen CT-Aufnahmen um fehlende Informationen, wie z.B. die Lage der Basilmembran, ergänzt werden können. Um die Fusion mit anderen Modalitäten einfach zu gestalten und die Verwendung in konventioneller Software zu gewährleisten, war es Ziel, aus den Schlifffbildern DICOM-Daten zu generieren.

2 Material und Methode

2.1 Probenvorbereitung

Das Präparationsverfahren basiert auf einem Kalteinbettverfahren mit Epoxidharz. Die zugeschnittenen, in einem Phosphatpuffer fixierten und in einer aufsteigenden Alkoholreihe dehydrierten Proben wurden mit Hilfe einer Silikongussform in gefärbtem Epoxidharz eingebettet (Abb. 1a), welches mit Hilfe einer Vakuumkammer entgast wurde, um sämtliche Gasblasen aus Kavitäten der Probe zu entfernen. Durch das Färben des Harzes mit undurchsichtigem Farbstoff wird sichergestellt, dass lediglich in der Schliffebene liegende Strukturen sichtbar sind (Abb. 1b), was die Identifizierung von Konturen bei der späteren Segmentierung erleichtert. Zur Erprobung des Präparationsverfahrens wurde zunächst ein Referenzkörper herangezogen, dessen bekannte geometrische Dimensionen eine Beurteilung der Rekonstruktionsgüte erlaubten (LEGO®Group, Billund, Dänemark). Im Folgenden wurde das Verfahren dann auf frisch entnommene, humane Mittel- und Innenohrpräparate angewendet.

Nach dem Aushärten wurden die Proben auf das Maß eines speziell angefertigten Handprobenhalters (Abschn. 2.2) abgedreht. Im Falle der Humanpräparate wurden zusätzliche Nuten auf dem Zylindermantel gefräst, um für die Beurteilung der Güte der Fusion der verschiedenen Datensätze über Referenzmarkierungen zu verfügen. Anschließend erfolgte der Scan der Proben mittels flächendetektorbasierter Volumen-CT (fd-VCT, GE Corporate R&D).

Für die Registrierung der einzelnen Schlifffbilder wurden mit Hilfe einer CNC-Maschine 3,2 mm durchmessende Bohrungen senkrecht zur Schliffebene eingebracht und diese mit kontraststarken Kunststoffstiften gefüllt (Abb. 1b). Die Lage der Bohrungen wurde mit Hilfe des jeweiligen fd-VCT Datensatzes geplant, um keine relevanten anatomischen Strukturen zu beschädigen. Für die Übertragung der Planungsdaten auf die Probe wurden die gefrästen Nuten her-

angezogen. Der bekannte Abstand der Registrierungsmarker diene später zudem zur Skalierung der Bilder.

2.2 Schliffpräparation

Um einen präzisen, äquidistanten Abstand der Schliffbilder zu realisieren, wurde ein spezieller Probenhalter entwickelt (Abb. 1b), welcher über ein Feingewinde die genaue Einstellung des gewünschten Schichtabtrages erlaubt. Dazu wird ein abriebsfester Hartkeramikring bezogen auf die Probenoberfläche derart verstellt, dass die gewünschte Probendicke diesen überragt und bei Druck auf die Schleifscheibe bis auf die Ebene des Keramikringes abgetragen wird. Im konkreten Fall wurden $100\ \mu\text{m}$ Schritte gewählt, was einen Kompromiss zwischen Rekonstruktionsgenauigkeit und zeitlichem Gesamtaufwand bei ca. 5 min Bearbeitungszeit pro Schichtbild darstellt. Geringere Schichtabstände sind technisch jedoch ohne weiteres möglich.

Nach jedem Schleifschritt wurde der tatsächliche Abtrag mit einem Wegsensor (Heidenhain Specto ST 3078, Dr. Johannes Heidenhain GmbH, Traunreut) mikrometergenau vermessen und dokumentiert. Die Bilddatenakquirierung erfolgte mittels Auflichtmikroskopie und 5-Megapixel Digitalkamera.

2.3 Bildregistrierung und DICOM-Generierung

Für die automatisierte Registrierung der Schichtbilder wurde in Matlab (The MathWorks, Inc., Natick, MA, USA) ein entsprechender Algorithmus realisiert, welcher nach geeigneter Bildvorverarbeitung mittels Hough-Transformation die

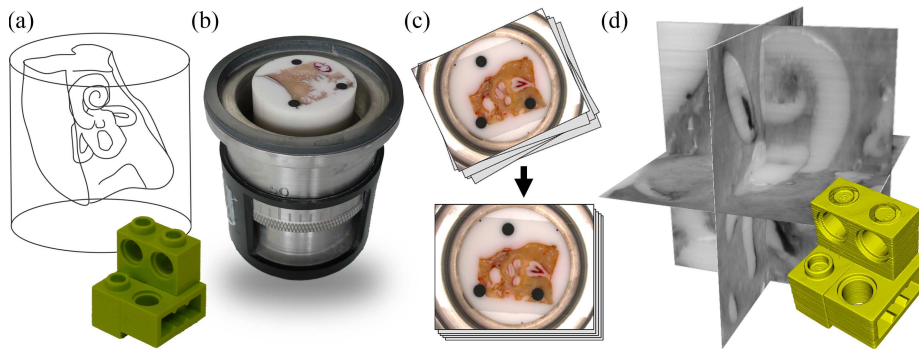


Abb. 1. (a) Schemazeichnung des eingebetteten und auf Mittel- und Innenohr beschnittenen Felsenbeinpräparat sowie der LEGO-Baugruppe. (b) Handprobenhalter mit eingespannter Probe. Sichtbar sind die drei eingebrachten Registrierungsmarker. (c) Schema der Registrierung des Bildstapels unter Verwendung der eingebrachten Registrierungsmarker. (d) 3D-Ansicht des als DICOM exportierten Schichtbilddatensatzes sowie die durch Segmentierung gewonnene 3D-Rekonstruktion des Referenzobjektes im STL-Format.

Registrierungsmarker detektiert (Abb. 1c) und entsprechend zum Basisbild ausgerichtet. Der so entstandene Bildstapel wurde anschließend auf die relevanten Strukturen beschnitten und diente als Grundlage für den DICOM-Export.

Der DICOM-Export erfolgte ebenfalls aus Matlab, um die für das DICOM-Format notwendigen Header-Einträge für jedes Bild spezifisch generieren zu können. Der Header-Einträge `pixel spacing` wurde aus dem tatsächlichen Abstand der Registrierungsmarker und dem entsprechenden Pixelabstand im Bild berechnet. Des Weiteren wurde die `slice thickness` zu 0,001 mm gewählt, um den zweidimensionalen Charakter der Schlifffbilder zu repräsentieren. Die `slice location` für jedes Schlifffbild errechnete sich als kumulativer Abstand aller bis dahin gemessenen Schichtabstände.

3 Ergebnisse

Der generierte DICOM-Datensatz konnte erfolgreich mit der kommerziellen Software PatXfer 5.2 eingelesen und in die Planungs- und Segmentierungssoftware iPlan 2.6 ENT (beide BrainLAB AG, Feldkirchen) übertragen werden. Damit stehen die histologischen Schlifffbilder als 3D-Datensatz gleichberechtigt zu anderen Modalitäten wie CT oder MRT zur Verfügung (Abb. 1d). Die Fusion dieser verschiedenen Bildgebungsverfahren ist damit möglich.

Die Segmentierung des Referenzobjektes erfolgte unter Verwendung eines schwellenwertbasierten Algorithmus. Dabei zeigte sich die zuverlässige Reproduktion der einstigen LEGO-Steine als nun virtuelles Modell (Abb. 1d). Nach der Segmentierung erfolgte der Export über die VVLink Schnittstelle ins STL-Format. Dies erlaubt die weitere Nutzung der Segmentierungsdaten in beliebigen CAD- oder FEM-Anwendungen. Konkret erfolgte der Import der STL-Datei in eine in C++ programmierte Vermessungssoftware – basierend auf dem Visualization Toolkit (Kitware, Clifton Park, NY, USA). Die dortige Vermessung der LEGO-Baugruppe diente zur Evaluierung der Zuverlässigkeit und Präzision der Rekonstruktion. So konnten rechte Winkel bzw. die Parallelität von Ebenen und Körperkanten mit einem mittleren Fehler von $0,3^\circ \pm 0,1^\circ$ erhalten werden. Die Länge von Körperkanten innerhalb der Schliffebene gelang mit einem mittleren Fehler von $2,0\% \pm 0,4\%$; Längen in Richtung des kritischen Schliffabtrages waren lediglich mit einem mittleren Fehler von $0,6\% \pm 0,3\%$ behaftet.

Bei den Felsenbeinpräparaten dienten die gefrästen Referenznuten zur Überprüfung der Rekonstruktionsgenauigkeit. Diese Strukturen sind sowohl im fdVCT-Datensatz als auch in den Schlifffbildern sichtbar, so dass nach der Fusion der beiden Datensätze deren Übereinstimmung vermessen werden konnte. Hier zeigten sich maximal Abweichungen von 0,2 mm, was in der Größenordnung der Auflösung der fdVCT-Daten liegt.

4 Diskussion

Die vorgestellte Methode erlaubt die hochgenaue 3D-Rekonstruktion anatomischer Strukturen aus histologischen Schlifffbildern. Diese können als Grundlage

zur Gewinnung von Anatomieatlanten für die modellbasierte Bildgebung dienen oder auch direkt als Basis für FEM-Untersuchungen genutzt werden.

Das gewählte Kalteinbettverfahren sichert den strukturellen Erhalt der wichtigen Weichgewebsstrukturen, wie z.B. der Basilarmembran, Stria vascularis oder Bänder und Muskel des Mittelohres. Darüber hinaus hat der Einsatz eines Schleifverfahrens im Gegensatz zum weitverbreiteten Schneiden mittels Mikrotom weitere Vorteile. So können harte Materialien wie Knochen direkt verarbeitet werden (wodurch der chemische Schritt der Dekalzifizierung entfällt) und metallische Implantate mit fixiert und ausgewertet werden. Die Einbettung in das harte Epoxidharz sichert zudem besser den Erhalt der geometrischen Abmessungen und Lagebeziehungen, als z. B. eine Einbettung in Paraffin, bei welcher durch die geringe mechanische Festigkeit der Einbettung, für eine exakte räumliche Rekonstruktion inakzeptable Gewebsverzerrungen zu verzeichnen sind.

In dieser Studie wurde erstmalig die Rekonstruktionsgenauigkeit anhand eines Referenzobjektes für eine derartige Präparationsmethode untersucht. Der mittlere Rekonstruktionsfehler von $0,6\% \pm 0,3\%$ zeigt deutlich, dass die Ausmessung des aktuellen Schichtabtrages und dessen spezifische Berücksichtigung beim Generieren der DICOM-Header ein zuverlässiges Verfahren ist.

Auch bei der Anwendung des Verfahrens auf Felsenbeinpräparate bestätigte sich die Zuverlässigkeit der Methodik. Da sich die Nuten im peripheren Bereich der Proben befinden und für den Fusionsalgorithmus nur die im Inneren liegenden Bereiche der eingebetteten Felsenbeine verwendet wurden, kann dort mit einem nochmals geringeren Fehler gerechnet werden. Im weiteren Verlauf der Arbeiten wird nun kontinuierlich die Stückzahl der aufbereiteten Proben erhöht und diese Daten als Grundlage für die Erstellung von hochaufgelösten 3D-Modellen des Mittel- und Innenohres einschließlich der Weichgewebsstrukturen verwendet.

Danksagung. Die Arbeiten wurden mit Mitteln des Bundesministeriums für Bildung und Forschung unter den Kennzeichen 01EZ-0832 /-0833 gefördert.

Literaturverzeichnis

1. Poznyakovskiy AA, Zahnert T, Kalaidzidis Y, et al. The creation of geometric three-dimensional models of the inner ear based on micro computer tomography data. *Hear Res.* 2008 Sep;243(1-2):95–104.
2. Sun Q, Chang KH, Dormer KJ, et al. An advanced computer-aided geometric modeling and fabrication method for human middle ear. *Med Eng Phys.* 2002 Nov;24(9):595–606.
3. Sørensen MS, Dobrzeniecki AB, Larsen P, et al. The visible ear: a digital image library of the temporal bone. *ORL J Otorhinolaryngol Relat Spec.* 2002;64(6):378–81.
4. Wang H, Northrop C, Burgess B, et al. Three-dimensional virtual model of the human temporal bone: a stand-alone, downloadable teaching tool. *Otol Neurotol.* 2006 Jun;27(4):452–57.
5. Santi PA, Rapson I, Voie A. Development of the mouse cochlea database (MCD). *Hear Res.* 2008 Sep;243(1-2):11–7.

Texture Analysis in Quantitative Osteoporosis Assessment Characterizing Micro-architecture in High Resolution Peripheral Quantitative Computed Tomography

Alexander Valentinitzsch¹, Janina Patsch¹, Dirk Mueller², Franz Kainberger¹,
Georg Langs^{3,1}

¹CIR Lab, Department of Radiology, Medical University of Vienna, Austria

²Philips GmbH Healthcare Division, Germany

³CSAIL, Massachusetts Institute of Technology, USA

`alexander.valentinitzsch@meduniwien.ac.at`

Abstract. High resolution peripheral quantitative quantitative computed tomography (HR-pQCT) permits in-vivo assessment of trabecular microstructure at an isotropic voxel size of $82\ \mu\text{m}$. The new imaging modality has potential in the differentiation of certain forms of osteoporosis based on the associated microstructural patterns. In this paper we propose an approach that assesses bone microarchitecture based on texture features extracted from the trabecular bone. The method is based on a three-dimensional gray level co-occurrence matrix descriptor, and a k-medians clustering in the feature space to indicate characteristic categories of texture. The distribution of the microarchitecture classes allows for a differentiation between osteoporotic and healthy subjects. We report initial results for the repeatability of the clustering method and the feasibility of the differentiation between healthy and osteoporotic bone for 6 subjects.

1 Introduction

Osteoporosis is a metabolic bone disease leading to an increased risk of fracture [1]. Low bone density and alterations in bone microarchitecture both contribute to increased fracture risk. Adequate diagnosis is the key to effective intervention. Osteoporosis remains clinically silent until the first fragility fractures occur [2]. The diagnosis and fracture risk assessment is currently based on bone mineral density (BMD) as observed by dual x-ray absorptiometry (DXA), which has certain limitations [3]. In particular the very specific, divergent bone metabolism associated with various pathophysiological subtypes of osteoporosis can not be differentiated by DXA alone. Thus, we consider texture analysis of high resolution peripheral quantitative computed tomography (HR-pQCT) data with an in-vivo resolution of $82\ \mu\text{m}$ as an important new approach with high potential in the differentiation of certain forms of osteoporosis.

In this paper we propose an approach that assesses bone micro-architecture based on texture features that are extracted from the spongiosa region in HR-pQCT volumes. A three-dimensional gray level co-occurrence matrix serves as a descriptor. A clustering in the resulting feature space indicates texture categories which appear repeatedly in the entire training set. Their distribution is used for the differentiation between osteoporotic and healthy subjects. It mimics the medical experts observation of distinct patterns in the regions, expected to correlate with different forms of the disease. Unlike current BMD assessment, this approach could indicate differentiation between different stages of osteoporosis or potentially identify sub-types of the disease. This differentiation could have a clinical impact by aiding treatment option decisions and leading to establishment of a novel biomarker which complements traditional BMD by DXA.

2 Methods

We first extract from training images local texture features that describe the micro-architecture quantitatively. These features are then clustered to derive categories of local texture. Unseen data is then processed by partitioning the entire volume into the respective classes identified by the training set. The method is divided into two main steps: (i) feature extraction with a three-dimensional gray level co-occurrence matrix, and (ii) k-medians clustering and classification of the extracted features of the trabecular bone.

2.1 Texture Analysis

To describe bone texture, quantitative feature extraction is performed using a three-dimensional gray-level co-occurrence matrix (3D GLCM) on the HR-pQCT volumes. The GLCM is a commonly used approach in medical imaging [4, 5]. In the volumetric case there are 13 different directions and we are using 4 voxel distances $d = \{1, 2, 4, 8\}$. To extract textural information from the GLCM, 12 different statistical Haralick texture measures [6] are computed for each direction. With linear scaling the original 12 bit-depth of the volume is normalized to 8 gray-levels. For every second point in the volume the local descriptor is calculated for a cubic neighborhood of $15 \times 15 \times 15$ voxels. On these overlapping subcubes a normalized symmetrical 3D GLCM with a re-quantization size of $Q = 8$ is performed. The number of gray levels Q determines the size of the GLCM and thus directly effects computational costs. The re-quantization to fewer gray levels in medical images has an advantage of reducing the noise-induced effects. For feature reduction and to avoid directional dependency, we calculate the angular mean and angular independent variance over all 13 matrices representing the directions.

2.2 K-medians Clustering to Obtain Texture Categories

After calculating the descriptor for each point in the spongiosa region of the HR-pQCT volume, we cluster feature vectors to capture sets of points with comparable characteristics. In this work we employ standard k-medians clustering. Instead of the squared Euclidian distance we utilize the Manhattan distance (L_1 norm). K-medians has been shown to be more sensitive to outliers for high dimensional features [7]. In our experiments it exhibited higher reproducibility and more feasible results than k-means. After a preliminary study and visual validation by a medical expert, we experimented with $4 \leq k \leq 6$.

2.3 Quantitative and Qualitative Validation

To validate the repeatability, we divided the available data into 2 disjoint sets: A, and B. We performed unsupervised learning of the clusters on one set (e.g., A), and applied the resulting classifier to the other set (e.g., B), and vice versa. The repeated measurements were evaluated by the Dice coefficient [8]. To describe the differentiation between the different groups a histogram of the cluster volume was performed, and the resulting distribution was compared to healthy and osteoporotic cases.

3 Results

We performed experiments on six HR-pQCT volumes of the distal radius having a size of $512 \times 512 \times 110$ with an isotropic resolution of $82 \mu\text{m}$. Three of the subjects were healthy volunteers and three were patients diagnosed as osteoporotic. For the k-medians clustering we compared 4 different parameter settings. Settings 1 and 2 use all 96 extracted texture features $d = \{1, 2, 4, 8\}$ with two different numbers of clusters $k = \{5, 6\}$. Settings 3 and 4 use a reduced feature space $d = \{4, 8\}$ with $k = \{5, 6\}$. We report initial quantitative results regarding 3 questions. Firstly, does the unsupervised differentiation of texture patterns in the bone result in a repeatable clustering? Put another way, can we expect the texture classes to represent true classes in the data? Second, can we use the frequency of the texture classes to differentiate between healthy and osteoporotic subjects? Finally, which texture classes are captured by the procedure, and do they correspond to texture classes intuitively described by medical doctors.

In Fig. 2(a) the results of the k-medians clustering are illustrated. The six texture categories corresponding to the feature clusters are feasible and distinguishable between each other. Minimal changes are visible according to the performance of the clustering. Dice similarity coefficients show also high levels of repeatability for our experiments with mean values of texture class overlap ranging from 0.906–0.990 and a standard deviation from 0.002–0.063. In Fig. 1 two examples of osteoporotic and healthy bone are depicted. The texture class images in Fig. 1(b, e) show differences in the distribution of texture classes, corresponding to the visible microstructural differences in the texture patterns

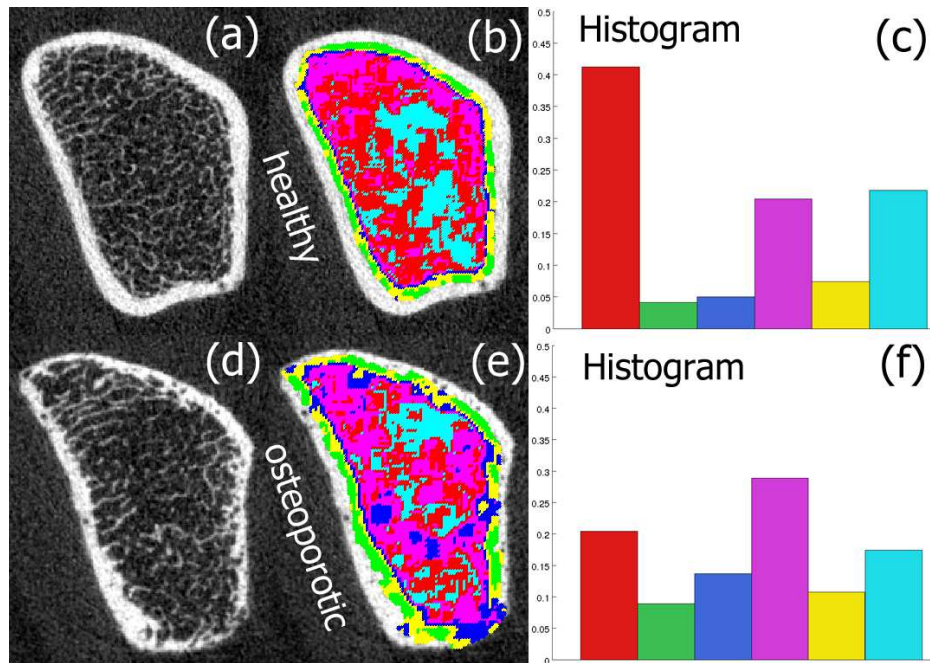


Fig. 1. Clustering comparison: healthy vs. osteoporotic bone microarchitecture: (a) is a healthy and (d) is a osteoporotic bone image slice, where k-means clustering with $k = 5$ was performed, shown in (b, e). In (c) the histogram of the k-means clustering result is illustrated, performed on the entire volume of the healthy bone (a). (f) represent the histogram of the osteoporotic case.

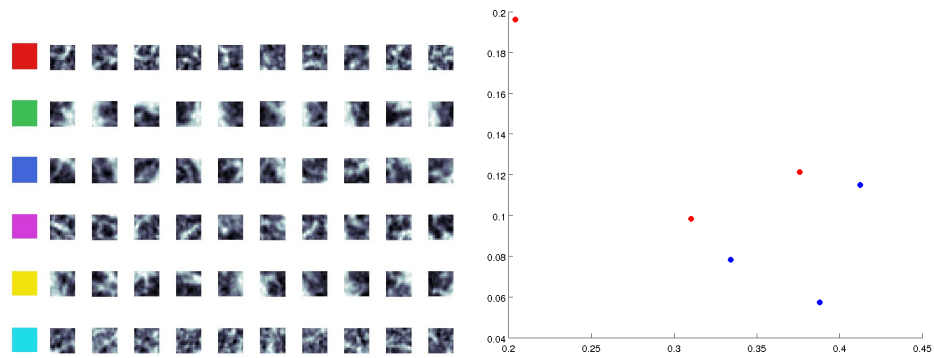


Fig. 2. Left: Prototype examples for six microarchitecture classes learned from the data. The patches show the ten nearest neighbors of each cluster centroid in the feature space; Right: Differentiation of osteoporotic (red dots) and healthy (blue dots) subject: x-axis is the cancellous buffer zone (red), y-axis is the cortical transition (yellow/green).

of Fig. 1(a, d). The Histograms vary between the osteoporotic and healthy subjects, as illustrated in Fig. 1(c, f). The cyan areas represent parts of the cancellous bone compartment which are poor in trabecular structure and rich in bone marrow. Trabecular core regions (magenta) are surrounded by trabecular buffer zones (red). Green and yellow indicate the outer border of the trabecular compartment and can be seen as a cortical transition zone with very thick trabeculae. Trabecular core and buffer zones are separated from the yellow/green cortical transition zone by a blue border region. The inverse relationship between the ratio of the cortical transition zone (yellow/green) and the cancellous buffer zone (red) and its correlation to osteoporotic versus healthy subjects indicate a potential differentiating power of the histogram features (Fig. 2(b)).

4 Conclusion

In this paper we investigate the differentiation between osteoporotic and healthy subjects according to their three-dimensional micro-architecture observed in HR-pQCT data. We learn spongiosa micro-architecture categories in an unsupervised manner and evaluate the repeatability of the resulting categories. Results indicate that the approach is applicable to define and classify micro-structural patterns of trabecular bone, and that the distribution of the micro-architecture classes exhibit a trend that allows for differentiating healthy and osteoporotic bone.

Acknowledgement. This work has been supported by the OeNB funded project Nr. 13468 (BIOBONE).

References

1. Bouillon R, Burckhardt P. Consensus development conference: diagnosis, prophylaxis, and treatment of osteoporosis. *Am J Med.* 1993;94:646–50.
2. Kazakia GJ, Majumdar S. New imaging technologies in the diagnosis of osteoporosis. *Rev Endocr Metab Disord.* 2006;7:67–74.
3. Kanis J, Kanis J. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: synopsis of a WHO report. *Osteoporos Int.* 1994;4(6):368–81.
4. Kovalev VA, Kruggel F, Gertz HJ, et al. Three-dimensional texture analysis of MRI brain datasets. *IEEE Trans Med Imaging.* 2001;20(5):424–33.
5. Herlidou-Meme S, Constans J, Carsin B, et al. MRI texture analysis on texture test objects, normal brain and intracranial tumors. *Magn Reson Imaging.* 2003;21(9):989–93.
6. Haralick RM, Shanmugam K, Dinstein IH. Textural features for image classification. *IEEE Trans Syst Man Cybern.* 1973;3(6):610–21.
7. Guha S, Mishra N, Motwani R, et al. Clustering data streams. In: *Proc Annu Symp Foundations Comp Sci*; 2000. p. 359–66.
8. Dice LR. Measures of the amount of ecologic association between species. *Ecology.* 1945; p. 297–302.

Lokale Analyse von Infarkttrandzonen in 3D-DE-MRT Bildsequenzen

Dennis Säring¹, Kai Müllerleile², Michael Groth³,
Gunnar Lund³, Heinz Handels²

¹Institut für Medizinische Informatik,
²Kardiologie mit Schwerpunkt Elektrophysiologie,
³Klinik für Diagnostische und Interventionelle Radiologie,
Universitätsklinikum Hamburg-Eppendorf
d.saering@uke.uni-hamburg.de

Kurzfassung. Bei Patienten mit Herzinfarkt kann es im weiteren Krankheitsverlauf zu lebensbedrohlichen ventrikulären Rhythmusstörungen kommen. Aktuelle medizinische Studien berichten, dass in MRT-Bilddaten heterogene Intensitäten im Infarkttrandbereich ein Hinweis auf mögliche Rhythmusstörungen sein können. In dieser Arbeit werden Methoden zur Extraktion von Parametern aus kontrastmittelgestützten delayed-enhancement MRT-Bilddaten vorgestellt, die mit hoher lokaler Auflösung eine quantitative Analyse des Infarkttrandbereiches anhand der Transmuralität, Hyperintensität und Heterogenität ermöglichen. Eine erste Evaluation der Verfahren wurden an 15 Datensätzen von Infarktpatienten durchgeführt. Die extrahierten Parameter wurden kombiniert visualisiert und durch medizinische Experten mit dem Wissen über den patientenspezifischen Krankheitsverlauf qualitativ bewertet. Die kombinierte Visualisierung wurde von den Kardiologen als hilfreich für eine schnelle Beurteilung des Infarktes eingeschätzt. In den vorliegenden Daten konnte gezeigt werden, dass die vorgestellten Parameter Hinweise auf mögliche Rhythmusstörungen liefern und somit die Unterstützung bei der Therapieplanung von Patienten mit akutem Infarkt verbessert werden kann.

1 Einleitung

Der Herzinfarkt ist eine der häufigsten Ursachen für frühzeitige Sterblichkeit in Deutschland. In der Zeit nach dem Infarkt kann es durch die entstandenen Narben im Herzmuskelgewebe zu lebensbedrohlichen ventrikulären Rhythmusstörungen kommen. Die Risikobeurteilung für Rhythmusstörungen erfolgt bisher auf Basis von wenigen quantitativen Parametern, wie einer stark reduzierten linksventrikulären Pumpfunktion (LVEF). Erste medizinische Fachpublikationen berichten, dass eine Koexistenz von vitalem und fibrotischem Myokard in der heterogenen Infarkttrandzone ein ideales Substrat für die Entstehung von ventrikulären Rhythmusstörungen bildet [1, 2]. Daher ist für klinische Studien neben der Lokalisation des Infarktes die Quantifizierung der Heterogenität von besonderem Interesse.

In den letzten Jahren wurden Ansätze zur Segmentierung von Infarktnarben [3] sowie die Quantifizierung der prozentualen Ausbreitung der Infarktnarbe, der sogenannten Transmuralität, auf Basis von 4D-Cine-MRT Bildsequenzen sowie kontrastmittelgestützte delayed-enhancement MRT Bilddaten (3D-DE-MRT) veröffentlicht [4]. Die Heterogenität eines Bildbereiches wird häufig durch Haralicksche Texturmerkmale, wie der local homogeneity, quantifiziert [5]. Grundlegend dafür ist die Cooccurrence-Matrix, welche die auftretenden Häufigkeiten der Grauwertkombinationen repräsentiert. Diese erfordert jedoch im zu untersuchenden Bildbereich eine hohe Anzahl an Pixeln, wodurch die Methode bei den vorliegenden Datensätzen nicht anwendbar ist. Lokale Form- und Funktionsanalysen des linken Ventrikels basieren häufig auf dem Ansatz der Centerline-Methode [6]. Dabei wird der komplette Herzmuskel auf Basis der myokardialen Konturen in gleichbreite Segmente unterteilt und segmenweise analysiert. Die Verwendung dieser Methode zur lokalen Analyse des Infarktgebietes sowie ein Verfahren zur Quantifizierung der Heterogenität in 3D-DE-MRT sind dem Autor nicht bekannt.

Ziel dieser Arbeit ist die Entwicklung neuer Methoden zur quantitativen Analyse des linken Ventrikels in 3D-DE-MRT mit hoher lokaler Auflösung. Hierbei sollen die für aktuelle klinische Studien interessanten Infarkt-Charakteristika, wie Transmuralität, Hyperintensität sowie Heterogenität in den Infarkttrandbereichen analysiert werden. Es soll untersucht werden, ob diese Parameter Hinweise zur Entstehung von ventrikulären Rhythmusstörungen liefern können.

2 Material und Methoden

2.1 Material

Für die Diagnose und Therapieentscheidung bei Patienten mit Herzinfarkt werden neben 4D-Cine-MRT auch häufig 3D-DE-MRT Bilddaten aufgenommen. Insbesondere die 3D-DE-MRT ist als nicht-invasive Referenzmethode zur Differenzierung zwischen vitalem und fibrotischem Myokard anerkannt [7]. Für die Entwicklung und Evaluierung der vorgestellten Methoden wurden MRT-Datensätze von 15 Patienten in der akuten Phase des Herzinfarkts verwendet. Jeder Datensatz besteht aus einer 3D-DE-MRT Bildsequenz. Die Aufnahme erfolgte EKG-getriggert in der enddiastolischen Phase des Herzrhythmus mittels T1-gewichteter turboFLASH (turbo fast low-angle shot) Sequenz. Die Auflösung der Datensätze beträgt 256×224 Pixel bei 7–12 Schichten pro Patient.

2.2 Unterteilung des Herzmuskels

Für die quantitative Analyse mit hoher lokaler Auflösung werden zunächst die myokardialen Konturen nach dem Verfahren von Säring et al. [4] semi-automatisch segmentiert. Anschließend wird automatisch auf Basis dieser Konturen in Anlehnung an die Centerline-Methode eine Unterteilung des Herzmuskels in 50 gleichbreite Segmente vorgenommen. Jedes Segment dieses 50-Segment-Modells repräsentiert somit 2% des Umfangs des gesamten Myokards (Abb. 1).

2.3 Segmentierung des Infarktes

Die Segmentierung des Infarktes erfolgt semi-automatisch. Hierzu werden zunächst fünf Regions of Interest (ROI) im gesunden Herzmuskelgewebe interaktiv festgelegt und der Mittelwert \bar{m} der Intensitäten sowie die Standardabweichung σ berechnet. Anschließend werden mittels Schwellwertverfahren alle Pixel innerhalb des Myokards mit einer Intensität von $I(\mathbf{x}) > \bar{m} + 2 \cdot \sigma$ als Infarkt definiert, wobei $I(\mathbf{x})$ der Intensitätswert an der Position (\mathbf{x}) im Bild angibt [8].

2.4 Extraktion der Parameter

Auf Basis des Segmentmodells und der Infarktsegmentierung werden die Parameter Transmuralität, Hyperintensität sowie Heterogenität voll automatisch bestimmt:

- Die *Transmuralität* T beschreibt die Größe und Lokalisation des Infarktes im Bezug zur lokalen Muskeldicke und wird prozentual angegeben. Hierzu wird für jedes Segment das Verhältnis von infarziertem und gesundem Muskelgewebe in Bezug auf die mittlere lokale Dicke des Segments berechnet. Eine Transmuralität von 100% repräsentiert einen Bereich der vollständig infarziert ist.
- Die *Hyperintensität* H kann direkt über die Unterschiede der Intensitäten vom betroffenen und gesundem Muskelgewebe berechnet werden und ermöglicht die Differenzierung zwischen vitalem und fibrotischem Myokard. Hierbei wird der für die Infarktsegmentierung berechnete Mittelwert \bar{m} als Referenzwert für die Intensität von gesundem Myokard definiert. Für alle myokardialen Voxel wird die prozentuale Abweichung

$$H(\mathbf{x}) = \left(\frac{I(\mathbf{x})}{\bar{m}} - 1 \right) \cdot 100 \quad (1)$$

vom Mittelwert berechnet. Da in den vorliegenden DE-MRT Bilddaten der Infarktbereich durch hohe Intensitäten dargestellt ist, sollte $H(\mathbf{x}) \gg 0$ auf fibrotisches Gewebe hinweisen.

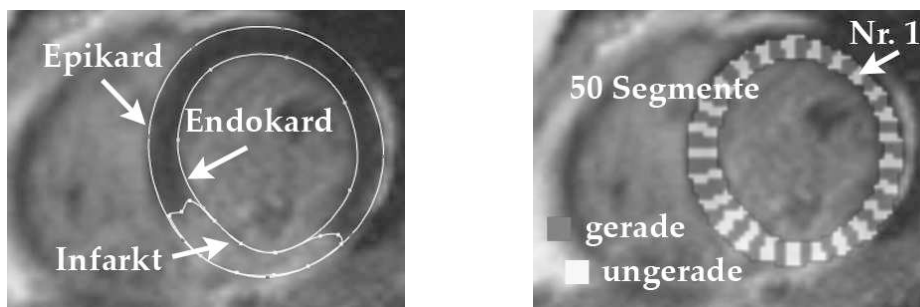
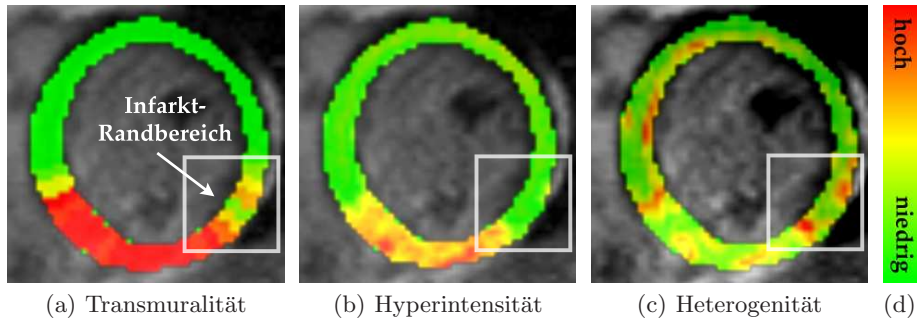


Abb. 1. Midventrikuläre 3D-DE-MRT Schicht mit der endo- und epikardialen sowie der Infarktkontur (links) und das generierte 50-Segment-Model (rechts).

Abb. 2. Farbcodierte Überlagerung der extrahierten Werte.



- Die *Heterogenität* M wird für jeden Pixel in einer lokalen 5-Umgebung über die Standardabweichung der Intensitäten $\sigma_{5 \times 5}$ berechnet, wobei nur die myokardialen Bildpunkte berücksichtigt werden. Ein kleiner Wert für $M(\mathbf{x}) = \sigma_{5 \times 5}(\mathbf{x})$ repräsentiert eine homogene 5×5 -Umgebung an der Bildposition \mathbf{x} . Die Visualisierung der Parameter erfolgt durch eine farbcodierte Überlagerung der quantitativen Werte auf dem Originalbild sowie durch eine segmentweise Mittelung aller Werte und die Kurvendarstellung dieser Mittelwerte für alle 50 Segmente. Kardiologen haben die Ergebnisse und deren Visualisierung mit dem aktuellen Wissen in Bezug auf bekannte ventrikuläre Rhythmusstörungen analysiert und qualitativ beurteilt.

3 Ergebnisse

Auf Basis der vorliegenden 15 Datensätze wurden mit den vorgestellten Methoden die klinisch relevanten Parameter T , H und M extrahiert und visualisiert.

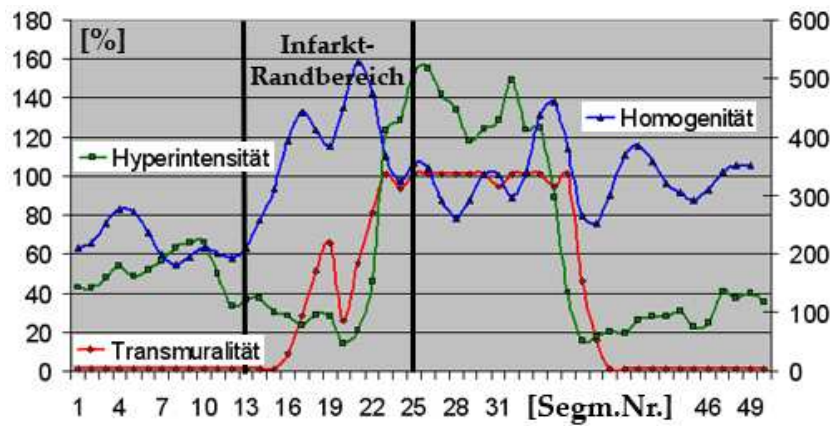


Abb. 3. 2D-Plot von T (rot), H (grün) und M (blau) für alle 50 Segmente.

Bei acht Patienten wurden stark heterogene Intensitäten im Infarkttrandbereich erkannt. Darunter waren alle sechs Patienten, bei denen im weiteren Krankheitsverlauf Rhythmusstörungen aufgetreten sind.

In Abb. 2 sind für einen Herzinfarktpatienten in midventrikulärer Schicht die extrahierten Parameter farcodiert überlagert dargestellt. In Abb. 3 sind für den gleichen Patienten die extrahierten Parameter für alle 50 Segmente farblich dargestellt. Als Grenzen des rechten Infarkttrandbereiches wurden die Segmente 14 und 25 manuell festgelegt, da dort die Transmuralität von 0 % auf 100 % ansteigt. In diesem Bereich ist eine stark erhöhte Heterogenität (blau) zu erkennen.

4 Diskussion und Ausblick

In dieser Arbeit wurden Verfahren vorgestellt, mit denen auf Basis von 3D-DE-MRT Bildsequenzen, Parameter zur quantitativen Beschreibung der Transmuralität, Hyperintensität und der Heterogenität extrahiert werden können. Um eine Analyse mit hoher lokaler Auflösung u.a. der Infarkttrandbereiche zu ermöglichen, wurde das Myokard in 50 Segmente unterteilt. Die extrahierten Parameter wurden segmentweise analysiert und visualisiert. Die kombinierte Darstellung von Transmuralität und Heterogenität ermöglicht eine schnelle visuelle Beurteilung des Infarktes. Eine erste qualitative Auswertung der 15 Datensätze hat ergeben, dass bei Patienten mit späterer ventrikulärer Rhythmusstörung heterogene Intensitäten in den Infarkttrandbereichen auftreten. Eine größere Studie soll dieses Ergebnis verifizieren. Weiterhin ist der Einsatz einer Support Vektor Maschine geplant, die eine Klassifizierung unter Verwendung aller Parameter ermöglicht. Es wird erwartet, dass dadurch die Risikobeurteilung einer Rhythmusstörung und somit die Unterstützung bei der Therapieplanung von Patienten mit akutem Infarkt weiter verbessert werden kann.

Literaturverzeichnis

1. Bogun FM, et al. Delayed-enhanced magnetic resonance imaging in nonischemic cardiomyopathy: utility for identifying the ventricular arrhythmia substrate. *J Am Coll Cardiol.* 2009;53(13):1138–45.
2. Yan AT, et al. Characterization of the peri-infarct zone by contrast-enhanced cardiac magnetic resonance imaging is a powerful predictor of post-myocardial infarction mortality. *Circulation.* 2006;114:32–9.
3. Kolipaka A, et al. Segmentation of non-viable myocardium in delayed enhancement magnetic resonance images. *Int J Cardiovasc Imaging.* 2005;21:303–11.
4. Säring D, et al. Computer-assisted analysis of 4D cardiac MR image sequences after myocardial infarction. *Method Inf Med.* 2006;45(4):377–83.
5. Handels H. *Medizinische Bildverarbeitung.* 2. Auflage. Vieweg + Teubner; 2009.
6. Sheehan FH, et al. Advantages and applications of the centerline method for characterizing regional ventricular function. *Circulation.* 1986;74(2):293–305.
7. Mahrholdt H, et al. Delayed enhancement cardiovascular magnetic resonance assessment of non-ischaemic cardiomyopathies. *Eur Heart J.* 2005;26:1461–74.
8. Kim RJ, et al. Relationship of MRI delayed contrast enhancement to irreversible injury, infarct age, and contractile function. *Circulation.* 1999;100:1992–2002.

Merkmale zur Beschreibung der Intensitätsvariation für die Klassifikation von Herdbefunden in Mammogrammen

Florian Wagner, Matthias Elter

Fraunhofer-Institut für Integrierte Schaltungen IIS, Erlangen
wagnerfn@iis.fraunhofer.de

Kurzfassung. Die computerassistierte Diagnose (CADx) für die Unterscheidung von Herdbefunden in Mammogrammen hat das Potential, Radiologen bei der Entscheidungsfindung zu unterstützen. Für ein von uns entwickeltes CADx-System haben wir eine Gruppe von neuen Merkmalen entwickelt, die auf einem Ansatz basieren, bei dem die Läsion in Abhängigkeit vom Abstand zum Rand der Segmentierung in verschiedene Regionen aufgeteilt wird. Für jede Region wird der Mittelwert der Grauwertintensität berechnet. Diese Mittelwerte dienen als Merkmal und als Eingabe für die Berechnung weiterer Merkmale, welche den Übergang zwischen den einzelnen Regionen und das Verhältnis zwischen den am weitesten entfernten Regionen beschreiben. Auf einer Testdatenbank mit 750 verschiedenen Läsionen erreichte die von einem Genetischen Algorithmus ermittelte optimale Merkmalsmenge auf 8 bit Bildern einen A_z -Wert von 0,72 und auf 12 bit Bildern einen A_z -Wert von 0,74. Die Ergebnisse zeigen, dass diese einfach zu berechnenden Merkmale gut für die Unterscheidung von benignen und malignen Herdbefunden geeignet sind.

1 Einleitung

Die Mammografie ist die aktuell wichtigste, für das Brustkrebs-Screening geeignete Methode und in der Mammadiagnostik unverzichtbar. Leider weist sie bei der Klassifikation von gut- und bösartigen Tumoren eine relative geringe Spezifität auf, sodass von den Patientinnen oftmals als unangenehm empfundene Folgeuntersuchungen durchgeführt werden müssen. Ein Beispiel für eine solche Untersuchung ist die Biopsie, bei der Gewebe zur pathologischen Untersuchung entnommen wird. Studien zeigen, dass nur 20 bis 30 Prozent der Biopsien bösartige Tumore identifizieren [1]. Für die Patientinnen bedeutet dies unnötige und vor allem unangenehme körperliche Eingriffe und für die Krankenkassen unnötige Kosten.

Um zur Reduktion von benignen Biopsien beizutragen, haben wir ein computerassistiertes Diagnose (CADx)-System zur automatischen Klassifikation von benignen und malignen Herdbefunden entwickelt. Das System berechnet für eine gegebene Läsion eine Vielzahl von Merkmalen (statistische Merkmale, Form-

und Texturmerkmale) und ermittelt anhand dieser die Wahrscheinlichkeit für die Bösartigkeit der vorliegenden Läsion. Es werden zudem Merkmale berechnet, die speziell bei Herdbefunden zu beobachtende Eigenschaften beschreiben.

Weisen Herdbefunde eine relativ hohe Dichte auf (d.h. sie erscheinen im Bild relativ hell), ist dies ein Indiz für einen malignen Befund. Transparenz ist hingegen eher ein Indikator für Gutartigkeit. Ein weiteres wichtiges Merkmal zur Charakterisierung ist die Umrandungsschärfe. Benigne Herdbefunde grenzen sich zumeist gut vom umgebenden Gewebe ab. Im Gegensatz dazu ist zu beobachten, dass maligne Läsionen oftmals fließend in das umgebende Gewebe übergehen. In dieser Arbeit wollen wir eine neue Gruppe von Merkmalen vorstellen, welche die Helligkeitsverteilung vom Zentrum des Tumors bis hin zum Rand beschreiben. Die Merkmalsgruppe ist in der Lage, die Dichte innerhalb des Tumors sowie den Übergang in das umgebende Gewebe zu beschreiben.

Varela et al. berechneten in [2] spezielle Merkmale für drei verschiedene Regionen des Herdbefundes. Unter anderem wurden Merkmale implementiert, die den Kontrast und die Umrandungsschärfe des Tumors charakterisieren. Die Klassifikationsleistung wurde anhand eines Datensatzes mit 981 Mammogrammen überprüft. Für jede Region wurde die optimale Merkmalskombination bestimmt. Die ROC-Analyse ergab für die innere Region einen A_z -Wert von 0.69, für die Randregion einen A_z -Wert von 0.76 und für die äußere einen A_z -Wert von 0.75. Die Kombination der Merkmale aller Regionen ergab einen A_z -Wert von 0.81.

Rangayyan et al. [3] präsentierten Merkmale, welche die Umrandungsschärfe und die Form von Herdbefunden beschreiben. Die Umrandungsschärfe eines Tumors wurde durch paarweise Pixeldifferenzen entlang der Normalen der Konturpixel charakterisiert. Auf einem Datensatz mit 54 Mammogrammen erzielte die beste Merkmalskombination bei der Klassifikation zwischen benignen und malignen Herdbefunden eine Genauigkeit von 94.4 Prozent. Das Merkmal zur Charakterisierung der Umrandungsschärfe zeichnete sich mit einer Genauigkeit von 92.6 Prozent als bestes Einzelmerkmal aus.

In [4] wurden Merkmale vorgestellt, die auf Grundlage eines polygonalen Modells der Kontur des Tumors berechnet werden. Nachdem der Rand des Herdbefundes als ein n -seitiges Vieleck modelliert wurde, berechnete man darauf Merkmale, welche die Konvexität bzw. Konkavität, die Wahrscheinlichkeit der Spikuliertheit und die Kompaktheit charakterisieren. Die Klassifikationsleistung aller Merkmale wurde auf einem Datensatz mit 54 Mammogrammen getestet. Das Merkmal für die Spikuliertheit eines Tumors wies den besten A_z -Wert von 0.82 auf. Die Merkmale für die Konvexität und Kompaktheit lieferten A_z -Werte von 0.75 bzw. 0.76. Die Kombination der drei Merkmale ergab einen A_z -Wert von 0.79.

Der Rest der Arbeit ist wie folgt gegliedert: In Kap. 2 stellen wir die für die Evaluierung verwendete Mammografie-Datenbank vor und beschreiben den Ablauf der Experimente sowie den Ansatz für die neue Merkmalsgruppe. Die Ergebnisse unserer Experimente werden in Kap. 3 vorgestellt und in Kap. 4 diskutiert.

2 Material und Methoden

Für die Auswertungen in dieser Arbeit wurde die öffentlich zugängliche Digital Database for Screening Mammography (DDSM) verwendet [5]. Berücksichtigt wurden alle Mammogramme von Fällen, die durch einen Lumisys Laserscanner digitalisiert wurden und mindestens einen Biopsie-geprüften Herdbefund enthielten. Für jeden Fall existiert eine Datei, welche die Annotation und die Grundwahrheit enthält. Für jedes Mammogramm ist eine manuell markierte Region von Interesse (ROI) gegeben, welche den Herdbefund enthält. Diese ROI wurde von einem Experten definiert, dient zur Lokalisation und als Eingabe für unser semi-automatisches Segmentierungsverfahren. Bilder die Artefakte oder Herdbefunde, die zu nah am Rand liegen enthalten, wurden ausgeschlossen, da unser Segmentierungsverfahren diese Fälle nicht behandeln kann. Letztendlich enthielt unsere Testdatenbank 750 Regionen (387 maligne und 363 benigne) von insgesamt 423 Patienten. Alle Bilder wurden mit einer Farbtiefe von 12 bit und einer Auflösung von $50 \mu\text{m}$ digitalisiert.

Die Grundidee des neu entwickelten Merkmalsextraktors ist die Beschreibung des Herdbefundes durch eine Vielzahl von konzentrischen Regionen, die in Abhängigkeit vom Abstand zu seinem Rand bzw. zum Rand der Segmentierung definiert werden. Um auch das umgebende Gewebe beschreiben zu können, wird die Segmentierung des Tumors in Abhängigkeit von der kleinsten Achse nach außen vergrößert. Durch die Anwendung der Chamfer-Abstands-Transformation (Details in [6]) auf die Segmentierung erhalten wir von jedem Pixel innerhalb der Segmentierung des Tumors den Abstand zur äußeren Begrenzung. Abhängig vom maximal berechneten Abstand $dist_{Max}$ teilen wir den Tumor in k konzentrische Regionen R auf. Jede Region R_i enthält alle Pixel $p(x, y)$ für die gilt

$$i \left(\frac{\text{maxValue}}{k} \right) \leq \text{dist}(p(x, y)) < (i + 1) \left(\frac{\text{maxValue}}{k} \right) + 1 \quad (1)$$

wobei $dist(p(x, y))$ den Abstand des Pixels $p(x, y)$ zum Rand des Tumors bezeichnet.

Für die k ermittelten Regionen werden dann die Grauwertistogramme H_R bestimmt und daraus der Mittelwert wie folgt berechnet

$$\text{Mean}_i = 1/||H_{R_i}|| \sum_{j=0}^{N-1} H_{R_i}(j)j, \quad i = 0, \dots, k - 1 \quad (2)$$

$||H_{R_i}||$ bezeichnet die Anzahl der Pixel in der Region i , N bezeichnet die Anzahl der Grauwerte und $H_{R_i}(j)$ steht für die Anzahl der Pixel mit dem Grauwert j in der Region i . Die k Mittelwerte der Regionen werden als Merkmale verwendet. Zusätzlich wird daraus ein Merkmal abgeleitet, welches die durchschnittliche Intensitätsveränderung zwischen den einzelnen Regionen charakterisiert

$$\text{IntensityChange} = 1/(k - 1) \sum_{i=1}^{k-1} |\text{Mean}_i - \text{Mean}_{i-1}| \quad (3)$$

Tabelle 1. Klassifikationsleistung (A_z) der vorgestellten Merkmale für $k = 10$.

Merkmal	8 bit	12 bit	Merkmal	8 bit	12 bit
Mean ₀	0,51	0,54	Mean ₈	0,59	0,63
Mean ₁	0,54	0,54	Mean ₉	0,64	0,66
Mean ₂	0,55	0,59	IntensityChange	0,62	0,63
Mean ₃	0,53	0,54	MaxRatio ₁	0,58	0,55
Mean ₄	0,57	0,59	MaxRatio ₂	0,59	0,56
Mean ₅	0,58	0,61	HistoDiff ₁	0,63	0,63
Mean ₆	0,61	0,65	HistoDiff ₂	0,63	0,62
Mean ₇	0,61	0,64			

Um den Kontrast des Tumors im Vergleich zum umgebenden Gewebe und damit auch die Umrandungsschärfe zu beschreiben, wurden die vier folgenden Merkmale konzipiert

$$\text{MaxRatio}_1 = \text{Mean}_{k-1} / \text{Mean}_0 \quad (4)$$

$$\text{MaxRatio}_2 = (\text{Mean}_{k-1} + \text{Mean}_{k-2}) / (\text{Mean}_0 + \text{Mean}_1) \quad (5)$$

$$\text{HistoDiff}_1 = \sum_{j=0}^{N-1} |H_{R_0}(j) / \|H_{R_0}\| - H_{R_{k-1}}(j) / \|H_{R_{k-1}}\|| \quad (6)$$

$$\text{HistoDiff}_2 = \sum_{j=0}^{N-1} |(H_{R_0}(j) / \|H_{R_0}\| + H_{R_1}(j) / \|H_{R_1}\|) - (H_{R_{k-1}}(j) / \|H_{R_{k-1}}\| + H_{R_{k-2}}(j) / \|H_{R_{k-2}}\|)| \quad (7)$$

Somit werden insgesamt $k + 5$ Merkmale berechnet, welche die Intensitätsverteilung innerhalb und außerhalb des Tumors sowie den Übergang in das umgebende Gewebe beschreiben.

3 Ergebnisse

Die Klassifikationsleistung der einzelnen Merkmale wurde durch die „leave-one-out“-Validation und „receiver operating characteristic“ (ROC)-Analyse ausgewertet. Die Fläche unter der ROC-Kurve A_z wurde als Referenz verwendet. Ein k-nächste Nachbarn (kNN)-Klassifikator wurde für den Klassifikationvorgang verwendet. Die vorgestellten Merkmale wurden sowohl für die 12 bit als auch für auf 8 bit herunter gerechnete Bilder für $k = 10$ berechnet. Um die gemeinsame Klassifikationsleistung aller Merkmale zu beschreiben, wurde durch einen genetischen Algorithmus eine möglichst optimale Merkmalsmenge ermittelt (sowohl für die 12 bit als auch für die 8 bit Merkmale).

Tabelle 1 zeigt die Klassifikationsleistungen für jedes einzelne Merkmal. Die A_z -Werte der jeweils in der optimalen Merkmalsmenge enthaltenen Merkmale

sind fett markiert. Für die auf 8 bit Bildern berechneten Merkmale erhält man eine optimale Gesamtklassifikationsleistung von 0,72 und für die auf 12 bit Bildern berechneten Merkmale einen A_z -Wert von 0,74.

4 Diskussion

Die in Tab. 1 dargestellten Ergebnisse zeigen, dass jedes einzelne Merkmal für die Unterscheidung von benignen und malignen Herdbefunden geeignet ist. Bei den einfachen Mittelwerten der Regionen (Mean_0 bis Mean_9) fällt auf, dass mit größerer Entfernung zum Rand der Segmentierung die Klassifikationsleistung steigt. Dies zeigt, dass die Grauwertintensität und somit auch die Dichte im Zentrum des Tumors charakteristischer für die Unterscheidung von gut- und bösartigen Läsionen ist als in den äußeren Bereichen. Außerdem ist zu beobachten, dass diese Merkmale, auf 12 bit Bildern berechnet, zu besseren Ergebnissen führen als das jeweilige, auf 8 bit Bildern berechnete Pendant. Somit spiegelt sich hier der Informationsverlust zwischen 12 bit und 8 bit in den Klassifikationsleistungen der einzelnen Merkmale wider.

Wir haben die Läsion in Abhängigkeit vom Abstand zum Rand der Segmentierung in verschiedene Regionen aufgeteilt und für jede Region den Mittelwert der Grauwertintensität berechnet. Diese Mittelwerte dienen als Merkmal und als Eingabe für die Berechnung weiterer Merkmale, welche den Übergang zwischen den einzelnen Regionen und das Verhältnis der am weitesten entfernten Regionen beschreiben. Die von einem Genetischen Algorithmus ermittelte optimale Merkmalsmenge erreichte auf 8 bit Bildern einen A_z -Wert von 0,72 bzw. einen A_z -Wert von 0,74 auf 12 bit Bildern berechnet. Die Ergebnisse zeigen, dass diese einfach zu berechnenden Merkmale gut für die Unterscheidung von benignen und malignen Herdbefunden geeignet sind. Es ist denkbar, dass eine Erweiterung der Merkmale, z.B. durch die Varianz der Grauwertverteilung in den einzelnen Regionen, zu einer Verbesserung der Gesamtklassifikationsleistung führen kann.

Literaturverzeichnis

1. Kopans DB. The positive predictive value of mammography. *AJR Am J Roentgenol.* 1992;158(3):521–6.
2. Varela C, Timp S, Karssemeijer N. Use of border information in the classification of mammographic masses. *Phys Med Biol.* 2006;51(2):425–41.
3. Rangayyan RM, El-Faramawy NM, Desautels JE, et al. Measures of acutance and shape for classification of breast tumors. *IEEE Trans Med Imaging.* 1997;16(6):799–810.
4. Rangayyan RM, Mudigonda NR, Desautels JE. Boundary modelling and shape analysis methods for classification of mammographic masses. *Med Biol Eng Comput.* 2000;38(5):487–96.
5. Heath M, Bowyer K, Kopans D, et al. The digital database for screening mammography. In: *Proc Int Works Digital Mammogr;* 2001. p. 212–8.
6. Butt MA, Maragos P. Optimum design of chamfer distance transforms. *IEEE Trans Image Process.* 1998;7:1477–84.

Plasmonen-unterstützte Mikroskopie zur Detektion von Viren

Frank Weichert¹, Marcel Gaspar¹, Alexander Zybin², Evgeny L. Gurevich²,
Alexander Görtz¹, Constantin Timm³, Heinrich Müller¹, Peter Marwedel³

¹Lehrstuhl für Graphische Systeme, Technische Universität Dortmund

²ISAS - Institut for Analytical Science, Dortmund

³Lehrstuhl für Eingebettete Systeme, Technische Universität Dortmund

`frank.weichert@tu-dortmund.de`

Kurzfassung. In Anbetracht zunehmend epidemisch auftretender viraler Infektionen ist eine effiziente und ubiquitär verfügbare Methode zur sicheren Virusdetektion hoch relevant. Mit der Plasmonen-unterstützten Mikroskopie steht hierzu eine neuartige Untersuchungsmethode bereit, die aber große Anforderungen an die Bildverarbeitung zur Differenzierung der Viren innerhalb der Bilddaten stellt. In dieser Arbeit wird hierzu ein erster erfolgversprechender Ansatz vorgestellt. Über bildbasierte Mustererkennung und Zeitreihenanalysen in Kombination mit Klassifikationsverfahren konnte sowohl die Differenzierung von Nanoobjekten als auch die Detektion von Virus-ähnlichen Partikeln nachgewiesen werden.

1 Einleitung

Im Kontext medizinisch-biologischer Fragestellung nehmen Biosensoren eine zunehmend hohe Bedeutung ein. Insbesondere der Anstieg epidemisch auftretender viraler Infektionen verlangt eine schnelle und zuverlässige Virusdetektion [1]. Mit der neuartigen Plasmonen-unterstützten Mikroskopie von Nanoobjekten (kurz: PAMONO-Technik, engl. **Plasmon assisted Microscopy of Nano-Size Objects**) [2] steht eine Methode zur Verfügung, welche die damit verbundenen Anforderungen in besonderer Weise unterstützt. Da bei der PAMONO-Technik auch einzelne Viren (allg. Nanoobjekte) in einer Flüssigkeit, z.B. Blut, nachgewiesen werden können, verbessert sich die Nachweisstärke signifikant gegenüber alternativen Verfahren und ermöglicht zusätzlich den Nachweis von Viren bereits ab einer Konzentration von 10^3 V/ml in der Probe. Damit ist die Nachweisfähigkeit vergleichbar zu etablierten Methoden wie beispielsweise ELISA oder PCR, bietet aber grundsätzliche Vorteile, da sie weniger empfindlich gegenüber Störungen ist. Die Zielsetzung des Projektes ist die Entwicklung einer portablen Echtzeit-Analysehardware, die es ermöglicht, einen Nachweis von Viren auch außerhalb spezieller Labore zu automatisieren. Die Schwierigkeiten liegen in den zu untersuchenden großen Datenmengen, dem relativ hohen Rauschanteil und den örtlich recht kleinen Ausprägungen der gesuchten Strukturen. Um dieses Ziel zu erreichen, wird in der vorliegenden Arbeit ein kombinierter Ansatz aus Zeitreihenanalyse [3] und bildbasierter Mustererkennung innerhalb einer mehrstufigen Klassifikationskaskade genutzt [4].

2 Material und Methoden

Im Folgenden wird der Versuchsaufbau des PAMONO-Sensors einleitend aufgezeigt, das aktuell umgesetzte Konzept der Datenaufbereitung und Merkmalsextraktion beschrieben und schließlich auf die Klassifikationen der gesuchten Strukturen eingegangen.

2.1 Plasmonen-unterstützte Mikroskopie

Die PAMONO-Technik beruht konzeptionell auf einer Ausnutzung der optisch-elektronischen Phänomene im Zusammenhang mit Oberflächen-Plasmonen und orientiert sich vom Versuchsaufbau her an der konventionellen Oberflächen-Plasmonen-Resonanz-Technik. Das Prinzip des PAMONO-Sensors (Abb. 1(a)) ist die Erkennung von markierungsfreien biomolekularen Bindungsreaktionen an einer Sensoroberfläche. Zur Detektion von Bindungsereignissen wird der Effekt ausgenutzt, dass polarisiertes Licht (Laser), welches gebündelt über ein Prisma auf eine Metallschicht trifft, reflektiert wird, dieses zu einer Anregung der Oberflächen-Plasmonen innerhalb der Metallschicht führt und eine Virusbindung die reflektierte Intensität verändert. Der Versuchsaufbau beinhaltet eine 12-Bit CCD-Kamera mit einer Auflösung von einem Megapixel (Pixelgröße $6.45 \mu\text{m} \times 6.45 \mu\text{m}$). Für die Aufnahme wurde bei einer auf $\sim 0.2 \text{ m}^2$ eingeschränkten Sensorfläche eine Frame-Rate von 40 fps (frames per second) an verfügbaren Bildern bei einer Auflösung von 100×1000 Pixeln erzielt. Für weitere Details zum Versuchsaufbau sei auf die Literatur verwiesen [2].

2.2 Merkmalsextraktion

Der initiale Schritt zur Merkmalsextraktion ist die Entfernung von irrelevanten systematischen Intensitätsvariationen und die Kompensierung des inhärenten Rauschens durch einen kombinierten Ansatz, der eine Hintergrundbereinigung mit gleitendem Referenzbild und eine zeitliche Mittelung aufeinanderfolgender

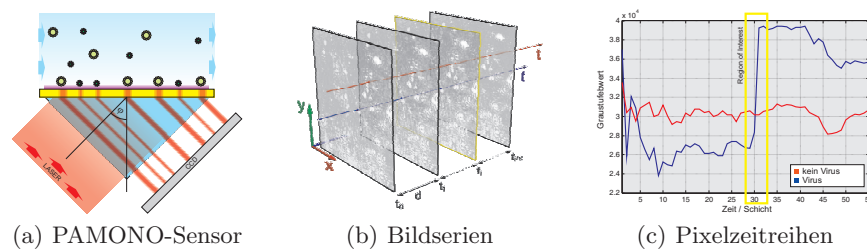


Abb. 1. Schematische Darstellung des Versuchsaufbaus. (a) Ausgehend vom PAMONO-Sensor wird (b) eine Bildfolge aufgenommen. (c) Vergleich des zeitlichen Intensitätsverlaufs an zwei beispielhaften Bildpositionen: die Anhaftung eines Virus an der Detektoroberfläche führt zu einem sprunghaften Anstieg.

Bilder beinhaltet. Die Abb. 1(b,c) vermitteln hierzu das Prinzip der Datenextraktion und Merkmalsgewinnung respektive das Konzept der Differenzierung zwischen Virus und Nicht-Virus. Aufgetragen über die Zeit resultiert bei dieser Vorgehensweise pro Pixel eine Signalkurve, dessen Charakteristik sich zwischen beiden Ausprägungen unterscheidet – signifikant ist der Sprung innerhalb der Viruskurve. Zum Zweck der Artefaktbehebung wird im Vorfeld der weitergehenden Analyse zunächst ein systematischer Intensitätstrend durch Subtraktion eines angepassten Polynoms entfernt. Desweiteren erfolgt zur signifikanten Reduktion von Fehlidentifikationen neben der zeitlichen auch die Ausnutzung der örtlichen Intensitätsänderung an einer Partikelposition. Dementsprechend werden durch nichtlineare Filterung der Bilder partikelähnliche Intensitätsverläufe mit zu geringer örtlicher Ausbreitung unterbunden.

2.3 Spektrale Zeitreihenanalyse

Zur Reduktion des Analyseaufwandes dient eine vorherige „schwache“ Partikelidentifizierung mittels einfachem Intensitätsschwellwert, sodass nur derart generierte Partikelkandidaten einer genaueren Analyse unterzogen werden. Im Hinblick auf die Klassifikation der Partikelkandidaten erfolgt eine Merkmalsextraktion über eine gewichtete Kombination aus Koeffizienten einer Zeit-Frequenz-Repräsentation und Differenzenquotienten mit zeitlich mittelndem Operatorfenster. Als Zeit-Frequenz-Repräsentation einer diskreten Funktion f dient dabei die S-Transformation $S_f(t_0, k)$ [5] gemäß

$$S_f(t_0, k) = \begin{cases} \frac{1}{N} \sum_{t=0}^{N-1} f(t) & \text{für } k = 0 \\ \sum_{m=0}^{N-1} F(m+k) \cdot e^{-\frac{2\pi^2 m^2}{k^2}} \cdot e^{i2\pi \frac{m}{N} t_0} & \text{für } k = 1, \dots, N-1 \end{cases} \quad (1)$$

mit $t_0 \in \{0, \dots, N-1\}$ und $k \in \{0, \dots, \lceil N/2 \rceil\}$. Dabei korreliert die S-Transformation die Zeitreihe mit einer Menge verschobener und frequenzmodulier-

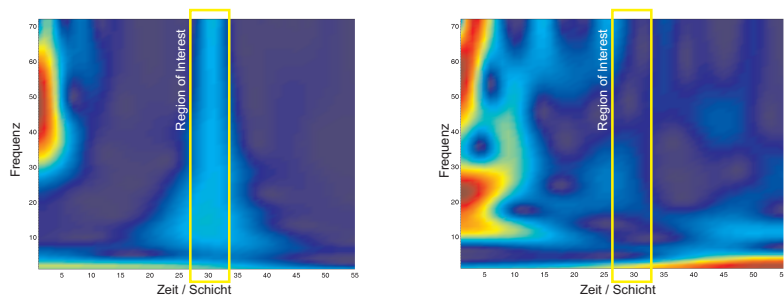


Abb. 2. Ausprägung von zwei Zeit-Frequenz-Darstellungen mittels der S-Transformation im Vergleich zwischen einer Virus- (links) und einer Nicht-Virus-Zeitreihe (rechts).

ter gaussförmiger Fensterfunktionen mit frequenzadaptiver Auflösung des Zeit-Frequenz-Bereiches. Abb. 2(a, b) zeigen die S-Transformationen der beiden Zeitreihen aus Abb. 1(c), welche mittels Gleichung 1 berechnet wurden. Die Abszisse zeigt die zeitliche Verschiebung t_0 des Analysefensters, und die Ordinate dessen zum Parameter k gehörige Frequenz ν_k in der Einheit Hz, wobei ν_k gegeben ist durch $\nu_k = k / (N \Delta T)$ mit $\Delta T = 1/40$ s.

2.4 Modellbasierte Klassifikation

Die extrahierten Partikelkandidaten werden modellbasiert gemäß ihrer Ähnlichkeit zum entsprechenden Klassenprototyp als Partikelpositionen bestätigt oder als partikelfreie Positionen verworfen. Bereiche aus benachbarten Partikelpixeln werden danach mittels Marching-Squares durch ein geschlossenes Polygon zusammengefasst. Ausgehend von Formmerkmalen findet abschließend eine Nachauswahl statt, um irrelevante Störungen mit partikelähnlichem zeitlichen Intensitätsverlauf anhand Eigenschaften ihrer örtlichen Ausdehnung zu verwerfen.

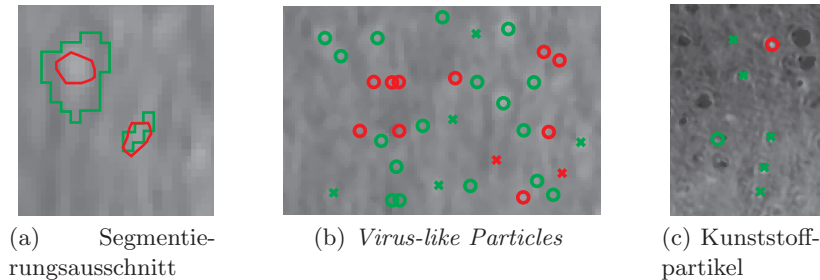
3 Ergebnisse

Zur Validierung der Klassifikation dienten Proben mit definierter Zusammensetzung – im Hinblick auf die Virus-Detektion ist die Verwendung von „*Virus-like Particles*“ (VLPs) zu nennen [6]. In Anlehnung an die ROC-Analyse wird die Nicht-/Fehl- bzw. korrekte Identifikation eines Partikels anhand von Überschneidungen der extrahierten Partikelpolygone mit manuell erstellten Referenzpolygone, welche nach Expertenwissen relevante Partikel markieren, ausgewertet (Abb. 4(a)). Überschneidungen extrahierter Polygone mit Referenzpolygone zählen dabei als True-Positive (TP), nicht geschnittene extrahierte Polygone als False-Positive (FP) und nicht geschnittene Referenzpolygone als False-Negative (FN). Der Ergebnistyp True-Negative (TN) ist auf diese Weise jedoch nicht eindeutig definierbar, da dies im Prinzip der Fall eines nicht erzeugten nicht geschnittenen Polygons wäre. Abb. 4(b) zeigt das Klassifikationsergebnis für die VLPs. In der quantitativen Analyse ergibt über drei ausgewertete Datensätze, bei einer Mittelung der Ergebnisse, mit 17 TPs und 3 FNs, eine TP-Rate von 0.85. Aufgrund des nicht definierbaren TN-Falles lässt sich keine FP-Rate definieren, aber eine durchschnittliche Anzahl von 9 FPs, welche noch als etwas zu hoch eingestuft wird. Zum Abgleich dieser Aussagen wurden zusätzlich Auswertungen mit Kunststoffpartikeln definierter Größe (80nm) durchgeführt, ebenfalls über drei Datensätze und Mittelung der Ergebnisse. Abb. 4(c) zeigt einen Ausschnitt der zugehörigen Segmentierung. In der quantitativen Analyse ergibt sich für die Kunststoffpartikel mit 17 TPs und 5 FNs eine TP-Rate von 0.77, bei durchschnittlich 3 FPs.

4 Diskussion

Ausgehend von der Motivation, ein sicheres und effizientes Verfahren zur Virusdetektion zur Verfügung zu stellen, wurde die PAMONO-Sensorik entwickelt.

Abb. 3. Partikelidentifikation: (a) Referenzpolygone sind rot (schwarz) und extrahier- te grün (grau). (b, c) Kreise markieren identifizierte Partikel, Kreuze Bereiche ohne erkennbare Partikel, grüne Markierungen korrekte und rote inkorrekte Klassifikationen der Pixelzugehörigkeit. Rote Kreise entsprechen Fehlalarmen, rote Kreuze unerkannten Partikeln, und grüne mittels Signalanalyse korrekt verworfenen Kandidaten.



Der neuartige Ansatz erlaubt insbesondere die explizite Detektion einzelner Viren auch in Proben mit geringer Konzentration. Da aber für eine ubiquitäre Nutzung dieser Technik eine automatische Auswertung unabdingbar ist, wurde innerhalb dieser Arbeit ein erster Schritt in diese Richtung unternommen. Durch die Kombination von Zeitreihenanalysen und merkmalsbasierter Klassifikation konnte aufgezeigt werden, dass eine Detektion von Nanopartikeln möglich ist. Insbesondere in Anbetracht unterschiedlicher Artefakte in den Datensätzen ist die quantitative Analyse positiv zu bewerten. Da zudem die Sensorik kontinuierlich weiter entwickelt wird, kann im Zuge der damit verbundenen Verbesserung der Bilddaten auch von einer deutlichen Reduktion der aktuell noch etwas hohen Anzahl von falsch positiven Strukturen ausgegangen werden. Neben der Verbesserung der Sensorik werden überwachte Lernverfahren zur automatisierten Parameteroptimierung und Generalisierung in den Ansatz integriert. Zudem wird an einer Modellierung der raum-zeitlichen Intensitätsvariation von Partikeln gearbeitet, um eine robustere und differenziertere Identifikation zu erlauben. Im Hinblick auf die angestrebte portable Analysehardware liegt ein übergeordnetes Ziel in der Überführung der Bildverarbeitungs- und Klassifikationsalgorithmen direkt auf Intelligente Kameras. Hierzu sind die Algorithmen weiter zu optimieren, damit eine Analyse derart großer Bild- und Datenmengen möglich ist.

Literaturverzeichnis

1. Mairhofer J, Roppert K, Ertl P. Microfluidic systems for pathogen sensing: A review. *Sensors*. 2009;9:4804–23.
2. Zybin A, inventor; Deutsche Patentanmeldung 10 2009 003 548.6. ; 2009.
3. Allen RL, Mills DW. *Signal Analysis*. Wiley-Interscience; 2004.
4. Niemann H. *Klassifikation von Mustern*. 2nd ed. Springer, Berlin; 2003.
5. Stockwell RG. *S-Transform analysis of gravity wave activity from a small scale network of airglow imagers [PhD Thesis]*. UWO; 1999. PhD Thesis.
6. Grgacic EV, Anderson DA. Virus-like particles: Passport to immune recognition. *Methods*. 2006;40(1):60–5.

Automatic Analysis of Live Cell Image Sequences to determine Temporal Mitotic Phenotypes

Nathalie Harder¹, Felipe Mora-Bermúdez², William J. Godinez¹,
Annelie Wünsche², Jan Ellenberg², Roland Eils¹, Karl Rohr¹

¹University of Heidelberg, BIOQUANT, IPMB, and DKFZ Heidelberg, Dept.
Bioinformatics and Functional Genomics, Biomedical Computer Vision Group

²European Molecular Biology Laboratory (EMBL) Heidelberg, Gene Expression and
Cell Biology/Biophysics Programmes

`n.harder@dkfz-heidelberg.de`

Abstract. Automated screening platforms allow biologists to acquire large amounts of image data with high information content. However, reliable automatic methods for analyzing this data are often not available. Here, we present an approach for detailed cell cycle analysis based on live cell fluorescence microscopy image sequences. Our approach comprises segmentation and tracking of dividing cell nuclei, and classifies cells into seven cell cycle phases as well as five abnormal morphological phenotypes. Moreover, we automatically quantify cell cycle phase durations and perform a statistical analysis to determine temporal phenotypes. Our approach was successfully applied to images from gene knockdown experiments and experiments treated with small molecule drugs.

1 Introduction

Understanding gene regulation of the cell cycle is of high common interest since errors in this process may lead to serious diseases such as cancer. High-content image-based screening is a powerful technology for gene function studies, and comprises automated microscopy as well as computational analysis to automatically extract information in an unbiased way. In screening experiments for cell cycle analysis usually live cell images of multiple cells are acquired. Multi-cell image sequences can be either analyzed in a population-based manner, i.e. features are determined for all cells and changes are studied for the whole population over time. However, subtle effects, such as cell cycle phase prolongations of certain cells, cannot be detected in this way. Therefore, single cell-based analysis has to be performed, which requires to track the individual cells throughout an image sequence. Based on tracking, the temporal evolution of single cells can be investigated, in particular, to study cell cycle phase progression. Previously, this has been done based on phase contrast ([1]) and fluorescence ([2, 3]) microscopy images. However, these studies distinguished only up to five cell cycle phases and did not consider abnormal cellular morphologies. Also, none of these studies determined cell cycle phase durations which is an essential readout for gene

function studies. In [4] seven normal phases were distinguished, but abnormal morphologies were not considered. Here, we present an approach to automatically determine morphological and temporal phenotypes on a single cell basis.

2 Materials and Methods

In this work, we analyze fluorescence microscopy 3D image sequences with three slices per time step, including multiple cell nuclei (Fig. 1, left). We developed the following image analysis approach: First, segmentation and tracking is performed based on maximum intensity projection (MIP) images. Next, image features are extracted based on original images and MIP images. Then, we perform classification of each nucleus at each time step, and finally process the obtained phase sequences and quantify phase durations using a finite state machine.

2.1 Segmentation and Tracking

For segmentation we developed a region-adaptive thresholding approach based on Otsu’s method. Local thresholds are computed in overlapping image regions and are applied to their non-overlapping center regions. For normal cell nuclei this method yields accurate results, however, for accurate segmentation of abnormal morphologies additional processing steps are necessary. In our application, abnormal morphologies are characterized by dim micronuclei attached to a bright main nucleus and small detached chromosomes. Proper segmentation of these structures is important to enable correct phenotype classification. For segmentation of dim micronuclei attached to a bright normal nucleus we apply the segmentation algorithm twice, where in the second run the bright normal nuclei are masked out. To merge single detached chromosomes to their corresponding main nucleus our approach automatically connects small objects (smaller than the smallest regular nuclei in anaphase) to the closest regular nucleus in their neighborhood by inserting a connecting line (Fig. 1, middle) .

For tracking dividing cells we developed a two-step approach. First, initial trajectories (*one-to-one* correspondences) are determined using a feature point tracking algorithm [5]. Second, mitosis events are detected and the respective trajectories are merged, establishing *one-to-many* correspondences. The detection of a mitosis event requires that (1) the daughter nuclei are smaller than the average nucleus, and (2) the Euclidean distance between daughter cells is below a threshold. If (1) and (2) are fulfilled, a measure for the likelihood of a mitosis event is computed, which is composed of three terms: The first term includes the ratio of the mean nucleus intensity and the intensities of the potential daughter nuclei, and yields values close to one if the daughter nuclei are very bright compared to the mean nucleus intensity. The second term considers the ratio of mother nucleus size and mean nucleus size, and yields values close to one if the mother nucleus is smaller than the average. The third term takes into account the analogous size ratio for the daughter nuclei, and consequently is close to one if both daughters are smaller than the average. The terms can be

individually weighted and sum up to a maximum value of one. If the likelihood measure is sufficiently high, the respective case is considered as mitosis (Fig. 1, right). Note that the number of possible daughter cells is not restricted to two, and thus, also abnormal splits into more than two daughter cells can be tracked.

2.2 Feature Extraction and Classification

We compute features for each cell nucleus based on the MIP images and based on single image slices of the original 3D images. In the latter case, the most informative slice is selected for each nucleus based on maximum total intensity. The reason for using the original image slices is that fine textures, which are important for the classification of certain phases (e.g., prophase), can be blurred in the MIP images leading to misclassifications. On the other hand, the selected slice not necessarily contains the entire object, e.g., detached chromosomes often were located in other slices. Consequently, we compute features related to texture, like Haralick texture features or wavelet features, based on the most informative slices, and features primarily related to object shape, e.g., size, circularity, or Zernike moments based on the MIP images. In addition, we use dynamic features representing the temporal changes of nuclei. These features are computed as differences of basic features (e.g., object size, mean and standard deviation of intensity, shape features) between subsequent frames. In total, we use 376 features. Our approach automatically classifies nuclei into 12 classes (inter-, pro-, prometa-, meta-, early ana-, late ana-, and telophase, abnormal early and late anaphase, abnormal telophase, abnormal interphase, and cell death) (Fig. 2, bottom, right) using support vector machines (SVMs) with a Gaussian radial basis function kernel. To obtain a more balanced data set we limited the number of interphase samples to 1000 per sequence, which yielded better results than using weighted SVMs.

2.3 Phase Sequence Analysis

To ensure consistency of the resulting phase sequences and to determine phase lengths we developed a finite state machine (FSM), which models cell cycle pro-

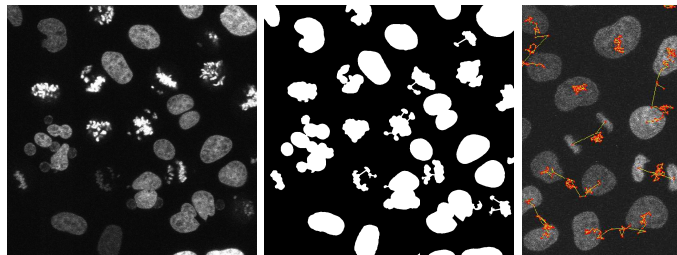


Fig. 1. MIP of a nocodazole treated experiment (left), segmentation with merged detached chromosomes (middle), enlarged tracking result with cell divisions (right).

gression and accepts only biologically plausible phase sequences. All 12 classes are represented by states of the FSM and all possible transitions between classes by state relations. Additionally, the FSM includes error states for all regular states that handle illegal phase transitions by correction or resetting. Note that using 12 instead of seven classes increases the model complexity significantly. The FSM processes phase sequences sequentially and each state includes a counter to determine phase durations.

3 Results

We applied our approach to 48 3D image sequences acquired on a confocal laser scanning microscope with an image acquisition interval of seven minutes and three slices per time step (1024×1024 pixels, 8 bit), and a total observation time of 15 h to 25 h. The imaged HeLa cells were fluorescently labeled with histone EGFP (visualizing the DNA), and treated with three different doses of the microtubule depolymerizing drug nocodazole (low, medium, and high). For each concentration we acquired six treated and six non-treated control sequences. Moreover, we performed siRNA knockdown experiments targeting the microtubule-associated gene *ch-TOG* and acquired six treated and six control sequences. Both treatments caused a delay in a prometaphase-like state, and afterwards chromosomal abnormalities such as lagging chromosomes, segregation defects, and appearance of micronuclei (Fig. 2, bottom, right).

The accuracy of our segmentation approach was evaluated based on four sequences where ground truth was determined manually. We quantified the occurrence of under- and oversegmentations and obtained an overall accuracy of 98.1%. The tracking accuracy was determined based on the same sequences, and we found that for a total number of about 16,900 matches, 40 mismatches occurred (mostly caused by segmentation errors), yielding an overall accuracy of 99.8%. Mitosis detection was evaluated based on 22 sequences, yielding an accuracy of 95.4% and a positive predictive value of 92.0%. The false positives were caused by abnormal morphologies such as detaching micronuclei. To determine

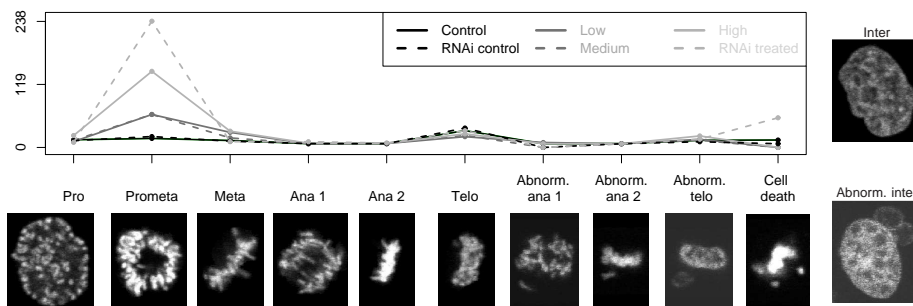


Fig. 2. Mean phase durations in minutes (y-axis) for the automatically analyzed image sequences, and example images for all 12 classes.

the classification accuracy we manually annotated 16 sequences from the nocodazole, and four sequences from the RNAi experiments. Using five-fold cross validation we yielded an overall classification accuracy of 93.9% for the nocodazole, and 94.7% for the RNAi data. Finally, we trained one SVM classifier with the annotated nocodazole and one with the annotated RNAi data, and applied them to the remaining, previously unseen test data. We determined the mitotic phase lengths based on the classification results using our finite state machine. To prove that our approach allows accurate determination of changes in mitotic progression, we analyzed the effects of nocodazole and siRNA treatment on the durations of mitotic phases. In particular, we tested whether prometaphase in perturbed cells was significantly longer than in controls, and compared the dose response of different nocodazole concentrations using Mann-Whitney tests ($\alpha=5\%$). We found a highly significant prometaphase prolongation for all nocodazole concentrations (Fig. 2). The high concentration showed a stronger effect of higher significance ($p=4.2\cdot 10^{-12}$) compared to the medium and low concentrations ($p=1.1\cdot 10^{-7}$ and $p=4.1\cdot 10^{-7}$). Cells treated with *ch-TOG* siRNA showed an even stronger increase of prometaphase duration (Fig. 2, $p=2.3\cdot 10^{-12}$).

4 Discussion

We presented an automatic approach for cell cycle analysis from live cell image sequences. Our approach allows accurate segmentation, tracking, and classification of normal as well as morphologically abnormal cell nuclei. We systematically evaluated the performance of the single steps of our approach based on real image sequences from different experiments. Our approach robustly quantifies the duration of mitotic phases and enables large-scale statistical analysis of phase durations to determine temporal phenotypes.

Acknowledgement. This work was supported by the EU projects MitoCheck and EuroDyna as well as the BMBF project NGFN-2.

References

1. Li K, Miller ED, Chen M, et al. Cell population tracking and lineage construction with spatiotemporal context. *Med Image Anal.* 2008;12(5):546–66.
2. Wang M, Zhou X, Li F, et al. Novel cell segmentation and online SVM for cell cycle phase identification in automated microscopy. *Bioinformatics.* 2008;24(1):94–101.
3. Padfield DR, Rittscher J, Thomas N, et al. Spatio-temporal cell cycle phase analysis using level sets and fast marching methods. *Med Image Anal.* 2009;13:143–55.
4. Harder N, Mora-Bermúdez F, Godinez WJ, et al. Determination of mitotic delays in 3D fluorescence microscopy images of human cells using an error-correcting finite state machine. In: *Proc. ISBI; 2007.* p. 1044–7.
5. Chetverikov D, Verestoy J. Tracking feature points: a new algorithm. In: *Proc ICIP; 1998.* p. 1436–8.

Model-Based Segmentation and Colocalization Quantification in 3D Microscopy Images

Stefan Wörz¹, Petra Sander², Martin Pfannmöller¹, Ralf J. Rieker^{3,4},
Stefan Joos², Gunhild Mechttersheimer³, Petra Boukamp⁵, Peter Lichter²,
Karl Rohr¹

¹Dept. Bioinformatics and Functional Genomics, Biomedical Computer Vision Group,
University of Heidelberg, BIOQUANT, IPMB, and DKFZ Heidelberg

²Dept. Molecular Genetics, DKFZ Heidelberg

³Dept. of Pathology, University Hospital, Heidelberg

⁴Institute of Pathology, Medical University of Innsbruck

⁵Dept. Genetics of Skin Carcinogenesis, DKFZ Heidelberg

s.woerz@dkfz.de

Abstract. We introduce a new model-based approach for automatic quantification of colocalizations in 3D microscopy images. The approach is based on different 3D parametric intensity models in conjunction with a model fitting scheme to quantify subcellular structures with high accuracy. The central idea is to determine colocalizations between different channels based on the estimated geometry of subcellular structures. Furthermore, we perform a statistical analysis to assess the significance of the determined colocalizations. We have successfully applied our approach to about 400 three-channel 3D microscopy images of soft-tissue tumors.

1 Introduction

Human telomeres are the physical ends of linear chromosomes that protect them from improper DNA metabolism. Telomere shortening within cell division plays an important role in aging and cancer. The activation of a telomere length maintenance mechanism (TMM) is thus essential for cellular immortalization and long-term tumor growth. Telomerase activity is the most common TMM. However, some tumors have been shown to utilize recombination-based TMMs that have been termed alternative lengthening of telomeres (ALT). Since the TMM can vary significantly in soft tissue sarcomas its characterization is crucial for developing new therapies. Promyelocytic leukemia (PML) bodies are sub-nuclear entities colocalizing with various proteins in the cell nucleus. ALT is indicated by preferential colocalization of PML bodies with telomeres as well as a heterogeneous telomere lengths with some very long telomeres [1].

This work aims at assessing ALT in human soft tissue sarcomas. Using a confocal laser scanning microscope, about 400 three-channel 3D images of soft-tissue sarcomas were acquired, visualizing telomere spots, PML bodies and DAPI-stained nuclei in the first, second, and third channel, respectively (Fig. 1).

The main task of image analysis is to automatically detect and classify colocalizations of telomeres and PML bodies as well as to detect and quantify very large telomere structures.

In previous work on quantifying the ALT mechanism, colocalizations were estimated manually or semi-automatically, which is subjective, error prone, and time consuming [1]. Automatic approaches for colocalization analysis can be classified into intensity correlation-based and object-based schemes. With intensity correlation-based schemes, global colocalization measures are computed, which directly correlate the intensity information of two channels [2]. Main disadvantages are the sensitivity w.r.t. image contrast and noise. In contrast, with object-based schemes a segmentation of the relevant image structures is obtained, which is used to determine colocalizations [3, 4]. However, so far typically spot-like structures have been considered using, e.g., threshold-based schemes [3] or the wavelet transform [4]. Moreover, these approaches usually only consider distances between spots to determine colocalizations [3, 4].

In this contribution, we introduce a new model-based approach for automatic quantification of colocalizations and very large telomere structures in three-channel 3D microscopy images. Our approach is based on 3D parametric intensity models in conjunction with a model fitting scheme to quantify telomeres and PML bodies with high accuracy. 3D intensity models have been previously used to segment subcellular structures [5, 6]. However, so far these approaches only used 3D Gaussian models and they have not been used for the detection of colocalizations. The central idea of our approach is to determine colocalizations of subcellular structures between different channels based on the estimated geometry. In contrast to previous approaches [3, 4], our approach not only allows to accurately detect colocalizations but also to differentiate between different types of colocalizations. Furthermore, we perform an automatic statistical analysis to assess the significance of the determined colocalizations.

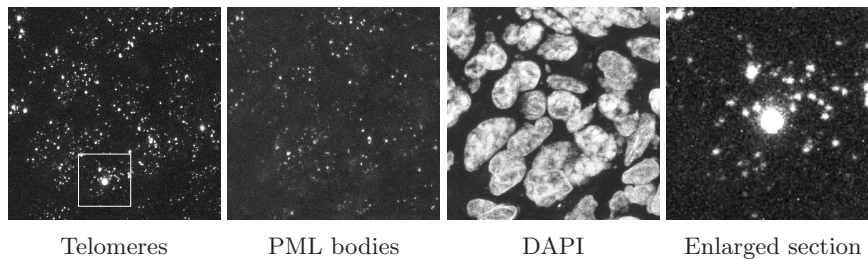


Fig. 1. Maximum intensity projections of 3D microscopy images of a tumor.

2 Materials and Methods

2.1 3D Model-Based Segmentation and Quantification

Telomeres and PML bodies appear as bright spots on a diffuse background (Fig. 1). To segment these structures, we propose a two-step approach. In the first step, we apply different image analysis operations to the 3D images to obtain coarse center positions of the spots [6]. In the second step, we quantify each detected spot based on model fitting. For PML bodies and small and medium sized telomeres the 3D intensity profile can be well represented by a 3D Gaussian

$$g_{\text{Gaussian3D}}(\mathbf{x}) = \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2} - \frac{z^2}{2\sigma_z^2}\right) \quad (1)$$

where $\mathbf{x} = (x, y, z)$ and $(\sigma_x, \sigma_y, \sigma_z)$ are the standard deviations. By including intensity levels a_0 (background) and a_1 (spot) as well as a 3D rigid transform \mathcal{R} with rotation $\boldsymbol{\alpha} = (\alpha, \beta, \gamma)$ and translation $\mathbf{x}_0 = (x_0, y_0, z_0)$, we obtain the 3D parametric intensity model $g_M(\mathbf{x}) = a_0 + (a_1 - a_0) g_{\text{Gaussian3D}}(\mathcal{R}(\mathbf{x}, \mathbf{x}_0, \boldsymbol{\alpha}))$. To fit the model to the image data, we use the non-linear optimization scheme of Levenberg/Marquardt. For each spot we obtain estimates of all model parameters with subvoxel resolution. However, for very large telomere structures (cf. the center spot in Fig. 1, right) the 3D intensity profile significantly deviates from a Gaussian shape. We propose to model these telomeres using a 3D superellipsoidal intensity model. Using $\Phi(x) = \int_{-\infty}^x (2\pi)^{-1/2} e^{-\xi^2/2} d\xi$, the model is given by

$$g_{\text{SuperEllipsoid3D}}(\mathbf{x}) = \Phi\left(\frac{\sqrt[3]{r_x r_y r_z}}{\sigma} \left(1 - \epsilon \sqrt{\left|\frac{x}{r_x}\right|^\epsilon + \left|\frac{y}{r_y}\right|^\epsilon + \left|\frac{z}{r_z}\right|^\epsilon}\right)\right) \quad (2)$$

where σ is the standard deviation of Gaussian image blur, (r_x, r_y, r_z) are the semi-axes of the superellipsoid, and ϵ describes its roundness. The final 3D parametric intensity model is obtained by including intensity parameters and a 3D rigid transform analogously to the Gaussian model. To segment the telomere channel, we propose a two-pass approach. In the first pass, the above described two-step approach with an adaptation to large spots is applied, i.e., using the superellipsoidal model. In the second pass, the smaller spots are segmented using the two-step approach where the regions of the large spots are masked out to prevent false spot candidates at large spots. To segment the cell nuclei in the third channel, we apply 3D Gaussian smoothing with subsequent Otsu thresholding.

2.2 3D Colocalization Quantification and Classification

To determine colocalizations between subcellular structures in different image channels, we propose the following scheme. The idea is to utilize the quantified geometry of the structures based on the fitting results of the parametric intensity models, i.e., the 3D position, 3D orientation, and 3D shape (specified by the standard deviations and semi-axes of the Gaussian and superellipsoidal model,

respectively). We define a colocalization as an (partial) overlap of two (or more) structures from different channels. Since, in contrast to previous approaches [3, 4], we determine a geometric representation of the structures with subvoxel resolution, we are able to differentiate between different types of colocalizations. In our application, we use in total six different types. The types c_1 to c_5 (Fig. 2) represent colocalizations between PML bodies (grey) and small/medium sized telomeres (black). In addition, c_6 denotes a colocalization of a PML body with a very large telomere signal, which is quantified using the superellipsoid. Overall, we count the number n_{c_i} of colocalizations of type c_i and the total number n_c .

2.3 Statistical Analysis of the Significance of Colocalizations

To assess the significance of the estimated number of colocalizations, we perform a statistical analysis given the null hypothesis that the observed colocalizations are a random result of the spot distribution. This analysis is based on repeated random placements of the segmented telomeres and PML bodies within the segmented volume of the DAPI channel. For each placement, the number of colocalizations is determined as described above. For all r random placements, Σ is the number of placements that yield a number of colocalizations greater than or equal to the experimental result. The empirical p-value is given by $p = \Sigma/r$, which can be compared to a certain level of significance α . Since we apply our approach to a large number of $n \approx 400$ images, we use a Bonferroni correction: given the experiment-wide false positive value π , the applied level of significance is $\alpha = \pi/n$. Thus, we can state for each colocalization type in each image whether the number of detected colocalizations is significant or not.

3 Results

We have applied our approach to about 400 three-channel 3D fluorescence microscopy images of human soft-tissue tumors. The images have a size of about $512 \times 512 \times 25$ voxels. For all images, we used the same parameter settings. Overall, we achieved quite good segmentation results. For example, Fig. 2 shows a maximum intensity projection of the segmentation results for the region marked

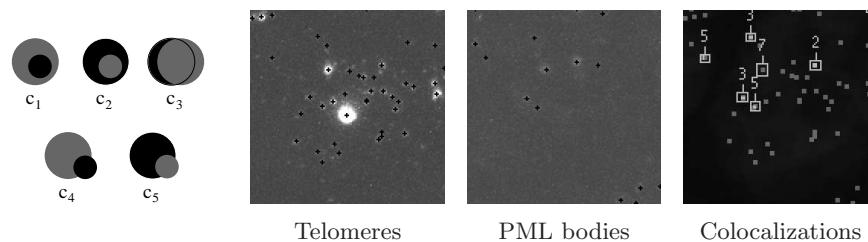


Fig. 2. Colocalization types c_1 to c_5 and results for the region shown in Fig. 1.

Table 1. Number of telomeres (n_{tel} , normal + very large) and PML bodies (n_{pml}) as well as total number n_c and numbers n_{c_1}, \dots, n_{c_6} of different colocalization types.

	n_{tel}	n_{pml}	n_c	n_{c_1}	n_{c_2}	n_{c_3}	n_{c_4}	n_{c_5}	n_{c_6}
Image 1	458+8	231	89	3	(1)	34	29	16	6
Image 2	479+5	181	22	(0)	(0)	(3)	(11)	(1)	7
Image 3	742+1	331	(16)	(0)	(0)	(0)	(16)	(0)	(0)

in Fig. 1 (left). Shown are the center positions of the segmented telomeres and PML bodies as well as the detected colocalizations together with their type.

For the image shown in Fig. 1 and for two additional images, Tab. 1 shows the number n_{tel} of segmented telomeres (normal + very large), the number n_{pml} of PML bodies, the total number n_c of colocalizations, as well as the numbers n_{c_1}, \dots, n_{c_6} of different types of colocalizations. Colocalization results which are not significant have been put in parenthesis. It can be seen that the number of colocalizations as well as the significance of these colocalizations is very different in the three images. Interestingly, whereas the number of segmented telomeres and PML bodies in the first two images is similar, the total number of colocalizations n_c differs by a factor of about 4. Still, the total number of colocalizations $n_c = 22$ of image 2 is significant. Considering individual types of colocalizations, it can be seen that the level of significance varies between different types. For example, for image 2 a number of $n_{c_4} = 11$ colocalizations is not significant, whereas a number of $n_{c_6} = 7$ colocalizations is significant. As a consequence, the statistical analysis of the significance of the colocalizations is important to allow an accurate analysis of the colocalization results.

4 Discussion

We introduced a new approach for automatic quantification of colocalizations and very large telomere structures in 3D microscopy images. The approach is based on 3D parametric intensity models in conjunction with a model fitting scheme to quantify telomeres and PML bodies with high accuracy. With our approach we determine colocalizations based on the estimated geometry (position and shape) of subcellular structures and differentiate between different types of colocalizations. Furthermore, we performed a statistical analysis to assess the significance of the colocalizations. We have successfully applied our approach to about 400 three-channel 3D fluorescent microscopy images of soft-tissue tumors. Future work is the analysis of the colocalization results w.r.t. biological meaning.

References

1. Fasching, CL et al. DNA damage induces alternative lengthening of telomeres (ALT)-associated promyelocytic leukemia bodies that preferentially associate with linear telomeric DNA. *Cancer Res.* 2007;67:7072–7.

2. Manders EMM, Verbeek FJ, Aten JA. Measurement of co-localization of objects in dual color confocal images. *J Microsc.* 1993;169:375–82.
3. Fay FS, Taneja KL, Shenoy S, et al. Quantitative digital analysis of diffuse and concentrated nuclear distributions of nascent transcripts, SC35 and Poly(A). *Exp Cell Res.* 1997;231:27–37.
4. Zhang B, Chenouard N, Olivo-Marin JC, et al. Statistical colocalization in biological imaging with false discovery control. In: *Proc ISBI*; 2008. p. 1327–30.
5. Thomann D, Rines DR, Sorger PK, et al. Automatic fluorescent tag detection in 3D with super-resolution: application to the analysis of chromosome movement. *J Microsc.* 2002;208:49–64.
6. Heinzer S, Wörz S, Kalla C, et al. A model for the self-organization of exit sites in the endoplasmic reticulum. *J Cell Sci.* 2008;121(1):55–64.

Prototyp eines Mammographie-CAD-Systems auf Basis der Cognition Network Technology

Ralf Schönmeier¹, Maria Athelougou¹, Harald Sittek², Peter Ellenberg¹,
Owen Feehan¹, Günter Schmidt¹, Gerd Binnig¹

¹Definiens AG, Trappentreustraße 1, 80339 München

²Diagnostisches Mammazentrum München, Feringastrasse 10b, 85774 München
rschoenmeyer@definiens.com

Kurzfassung. In dieser Arbeit wird ein neues Assistenzsystem für die computerunterstützte Detektion und Diagnose von Herdbefunden in der Mammographie vorgestellt, das mit Hilfe der Cognition Network Technology (CNT) eine bildbasierte Ähnlichkeitssuche in einer Referenzdatenbank ermöglicht. Dabei wird vorverarbeitend vollautomatisch eine mehrskalige Segmentierung und Klassifikation auffälliger Objekte durchgeführt, deren Ergebnisse zusammengefasst in der Datenbank zur Verfügung stehen. Die Ähnlichkeitssuche lässt sich interaktiv konfigurieren und der Prototyp ist offen für Erweiterungen und die Einbindung von Daten weiterer Modalitäten.

1 Einleitung

Die Früherkennung von Brustkrebs anhand von Mammographiedaten stellt eine wichtige und anspruchsvolle Aufgabe für den Radiologen dar. Hauptanhaltspunkte, ob ggf. eine gut- oder bösartige Brustläsion vorliegt, liefern typischerweise verschiedene Erscheinungsformen von Kalzifikationen und Verdichtungen, den sogenannten Herdbefunden. Da deren morphologische Variabilität sehr groß ist, lassen sich sichere Aussagen über Gut- oder Bösartigkeit oftmals nicht alleine aus der Bildgebung ableiten und erst eine Biopsie liefert im Zweifel eine sichere Diagnose. Die bereits verfügbaren Systeme zur computerunterstützten Detektion und Diagnose (CAD) von Brustläsionen basieren hauptsächlich auf Bildanalysealgorithmen und helfen dem Radiologen auffällige Gebiete in Bildern nicht zu übersehen [1, 2]. Unser Ansatz verfolgt die Strategie, zu einem gegebenen Patientenfall automatisiert in einer Referenzdatenbank relevante ähnliche Fälle zu finden und anhand zugehöriger Referenzbefunde die Qualität von Diagnosen zu steigern. Dazu wurde die CNT dahingehend erweitert, nicht nur wie bisher Bilddaten [3, 4], sondern auch Daten, die in Tabellenform vorliegen und auch textwertige Einträge aufweisen können, zu verarbeiten. Dies ermöglicht die Erstellung eines Datenbanksystems, das die Ergebnisse von bildanalytischen Untersuchungen und anderen Daten, wie z.B. Patienten-Meta-Daten, enthält. Mit CNT werden dann die Beziehungen dieser Einträge heterogenen Ursprungs für eine kontextgetriebene Analyse genutzt.

2 Material und Methoden

Der vorliegende CAD-Prototyp wurde auf Basis der in [5] vorgestellten Mammographie-Datenbank mit detaillierten Annotationen und bioptisch abgesicherten Befunden entwickelt. Als Entwicklungsplattform dient die CNT-Software *Definiens XD 1.2.0* [6, 7]. Um der großen Variabilität von Herdbefunden Rechnung zu tragen und für die weitere Verarbeitung adäquate Bildobjekte zu liefern, kommen bei der vorverarbeitenden Bildanalyse der einzelnen Mammogramme unterschiedliche Segmentierungsverfahren zum Einsatz. Daraus gewonnene Merkmalsvektoren werden in einer Datenbank zusammengefasst. Diese direkten Objekteigenschaften (wie z.B. Fläche und Helligkeit) stellen zusammen mit daraus abgeleiteten Kontexteigenschaften (z.B. Distanz zur ebenfalls automatisch detektierten Mamilla) die Entscheidungsgrundlage für einen Klassifikator zur Auffindung von auffälligen Verdichtungen dar. Neben der vollautomatischen Segmentierung und Klassifikation wird ein Konzept umgesetzt, mit dem sich ein Maß für *Ähnlichkeit* konfigurieren lässt, das dem Benutzer zu einem gegebenen Mammographie-Fall auf Knopfdruck *ähnliche* Fälle liefert. Im Folgenden werden die einzelnen Programmmodule etwas genauer skizziert.

2.1 Vorverarbeitung

Zunächst erfolgt eine Normalisierung der Grauwerte aller Eingangsbilddaten, um die unterschiedlichen Gegebenheiten verschiedener Gerätehersteller zu berücksichtigen. Es schließt sich eine Segmentierung der Brustkontur, die sich im jeweiligen Mammogramm als Kante darstellt und das Bild in Vorder- und Hintergrund unterteilt, an. Auf ihr erfolgt eine Abschätzung für die Position der Mamilla als dem Punkt mit dem größten Abstand zu einer Basislinie, die in cranio-caudalen (CC) Ansichten der rechten oder linken Brust mit dem rechten bzw. linken Bildrand zusammenfällt. In medio-lateralen (ML) Ansichten ist diese Basislinie um -15° bzw. 15° gedreht, was sich empirisch für eine ausreichend robuste Erkennung auf allen zur Verfügung stehenden Mammogrammen bewährt hat. Wenn in ML-Ansichten ein Teil des Pektoralmuskels sichtbar ist, so zeichnet sich dieser in der Regel als eine ausgeprägte Helligkeitsstufe aus. Diese kann mit Hilfe eines Kantenfilters auf einem stark geglätteten Bild detektiert werden. Die Abschätzung des Pektoralmuskels wird nach einer Plausibilitätsprüfung durch eine lineare Extrapolation der Hauptrichtung dieser Kante sowie dem Maskieren des Bereichs links (in linken Brustansichten) bzw. rechts (in rechten Brustansichten) davon vervollständigt. Für die Gewinnung von Kandidatenobjekten von potentiellen Herdbefunden werden fünf unterschiedliche Segmentierungen durchgeführt: ein Verfahren, das ausgehend von Startregionen mittels morphologischer Operationen auf relativ kontrastreiche Objekte abzielt, wird für vier verschiedene Größenskalen angewandt. Ein alternatives Verfahren segmentiert vorwiegend Objekte, die sich im Mammogramm recht kontrastarm abzeichnen und in einer relativ homogenen Umgebung befinden.

2.2 Tabellenverarbeitung

Die exportierten Daten der Bildanalyse werden in Tabellenform (CSV-Format) zusammen mit Patienten-Metadaten (wie z.B. Alter) als relationales Datenmodell in die Definiens Software geladen, bei dem zugehörige Daten miteinander „verlinkt“ sind. Dadurch stehen sämtliche Informationen in einem einheitlichen Zugriffsmodell zur Verfügung und es lassen sich auch erweiterte Merkmale formulieren, die Beziehungen zwischen Objekten unterschiedlicher Bilder ausnutzen. Diese Merkmale liefern z.B. ein Maß dafür, ob Kandidatenobjekte für eine Verdichtung im selben Abstand zur Mamilla sowohl in der CC- als auch ML-Aufnahme der selben Brust auftreten. Die Klassifikation auffälliger Herdbefunde erfolgt mittels eines hierarchisch aufgebauten Fuzzy-Klassifikators, der sich im Wesentlichen aus zehn Unterklassifikatoren für unterschiedliche Merkmalsprofile zusammensetzt.

2.3 Ähnlichkeitssuche

Die Suche nach ähnlichen Fällen zu einem gegebenen Fall basiert auf einem Maß für Patientenähnlichkeit, das sich hierarchisch aus Herdbefund-, Brustansicht- und Patienten-Metadaten-Ähnlichkeit zusammensetzt. Es werden jeweils

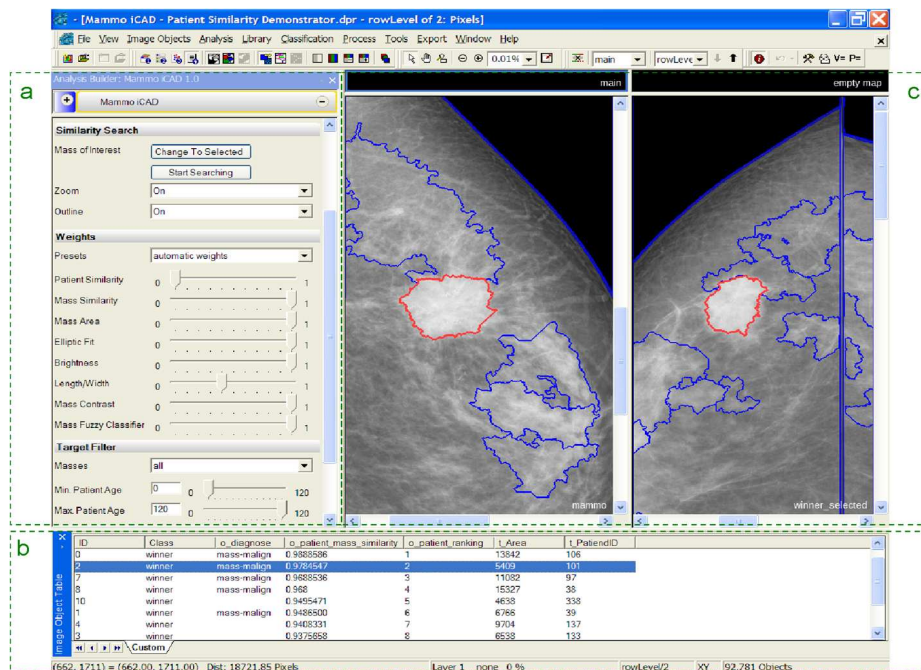


Abb. 1. Benutzerschnittstelle für die Ähnlichkeitssuche mit der Möglichkeit, die Suchanfrage zu konfigurieren (a), die Ergebnisse auszuwählen (b) und anzuzeigen (c).

fuzzybewertete Differenzen zwischen Merkmalswerten des aktuellen Anfragefalls mit allen Vergleichsfällen berechnet. Bei der Herdbefund-Ähnlichkeit kommt der zuvor beschriebene Klassifikator zum Einsatz, der eine auffällige Verdichtung vorschlägt. Diese wird dann für die Ähnlichkeitssuche mit den Kandidatenobjekten aller anderen Patienten der Datenbank herangezogen. Die Suchanfrage lässt sich interaktiv mittels Gewichtsvariablen konfigurieren, um den Einfluss einzelner Merkmale regeln zu können (Abb. 1 a). Als Ergebnis der Suche werden die für die aktuelle Konfiguration 10 ähnlichsten Fälle aufgelistet (Abb. 1 b) und die zugehörigen Bilddaten per Mausklick angezeigt (Abb. 1 c). Da die Befunde der Referenzfälle in der Datenbank ebenfalls vorliegen, erfolgt mit deren Hilfe ein automatisierter Diagnosevorschlag, der auf einem Mehrheitsvoting beruht.

3 Ergebnisse

Die Entwicklungsdatenbank, die auf sämtlichen Bildern aus [5] basiert, enthält ca. 58.000 Herdbefund-Kandidatenobjekte von 303 Patienten. Davon werden 314 Objekte, die mit Annotationen übereinstimmen, vom Klassifikator detektiert. Diese verteilen sich auf 90 unterschiedliche Patienten von den insgesamt 112 mit auffälligen Herdbefunden, was einer Trefferquote von 80% entspricht. Der Klassifikator liefert im Schnitt 0.66 falsch-positive Objekte pro Bild (677 Objekte in 1024 Brustansichten). In einer Testdatenbank, die ca. 34.000 Kandidatenobjekte für Herdbefunde von 120 neuen Patientendatensätzen enthält, detektiert der identische Klassifikator für 7 von den 15 ausschließlich benign annotierten Patienten richtig erkannte Herdbefunde. Bei einer Anzahl von 541 falsch-positiven Objekten aus 480 Bildern entspricht dies durchschnittlich 1.13 falsch-positiven pro Bild. Eine Ähnlichkeitssuche dauert rund 10 Sekunden für den Vergleich mit allen insgesamt 423 Patientenfällen und skaliert linear. Für den Test der Diagnose wurde für jeden dieser Fälle eine Ähnlichkeitssuche mit allen anderen Patienten durchgeführt, um nur anhand deren Referenzbefunde den Diagnosevorschlag zu ermitteln und dann mit dem eigenen Referenzbefund zu vergleichen. Es wurden zwei Szenarien getestet: ein Drei-Klassenproblem, bei dem für die Diagnose die Klassen *malign*, *benign* und *unauffällig* zur Verfügung standen (Tab. 1(a)) und einmal als Zwei-Klassenproblem mit einer Beschränkung des Suchraums auf nur *maligne* und *benigne* Vergleichsobjekte (Tab. 1(b)).

4 Diskussion

Die quantitative Leistung der Detektion von Herdbefunden erreicht das Niveau bestehender Verfahren [8], obwohl unser System auf Grundlage einer vergleichsweise großen Datenbasis auch die Erzeugung der ROIs (*regions of interest*) vollautomatisch leistet. Der Benutzer hat die Möglichkeit, diesen Vorschlag für die Ähnlichkeitssuche zu überstimmen und diese Information auch zu speichern. Dadurch „lernt“ die Referenzdatenbank und gewinnt an Wert für künftige Suchanfragen. Die Ähnlichkeitssuche liefert relevante Vergleichsfälle und erlaubt es dem Benutzer somit, die Referenzdatenbank zu durchsuchen, wie es mit manueller

Tabelle 1. Konfusionsmatrix für das (links) Drei- und (rechts) Zwei-Klassenproblem mit den relativen Anteilen der Klassifikation basierend auf allen 423 Patientendatensätzen der Datenbank bzw. von den insgesamt 97 (66 malignen und 31 benignen) automatisch detektierten Patienten mit auffälligen Herdbefunden.

Diagnose- vorschlag	Referenzdiagnose			Diagnose- vorschlag	Referenzdiagnose	
	malign	benign	unauffällig		malign	benign
malign	36%	6%	4%	malign	76%	25%
benign	0%	0%	0%	benign	25%	75%
unauffällig	64%	94%	96%			

Durchsicht nicht zu leisten ist. Für eine Verbesserung der Ergebnisse des automatischen Diagnosevorschlags wäre es wünschenswert, wenn eine umfangreichere Referenzdatenbank zur Verfügung stünde, um überhaupt adäquat ähnliche Fälle für die große Variabilität von Herdbefunden zu enthalten. Bereits während der Entwicklungszeit stellte der Prototyp eine leistungsfähige Umgebung zum Entwickeln, Testen und Evaluieren von Bildanalysen auf Basis vieler Patientendatensätze dar. Die Konzepte, die dabei erarbeitet wurden, sind nicht nur auf die Anwendung für die Mammographie beschränkt und weisen durch ihre Erweiterbarkeit den Weg zu CAD-Systemen, die auf multimodaler Bild- und Datenanalyse beruhen und unserer Meinung nach künftig einen großen Stellenwert einnehmen werden.

Danksagung. Diese Arbeit wurde von der Bayerischen Forschungsförderung im Rahmen des Projekts Mammo-iCAD gefördert.

Literaturverzeichnis

1. Destounis S, et al. Special session on breast CAD. *Int J CARS*. 2007;2(Suppl 1):S330–S350.
2. Doi K. Special session on breast CAD. *Int J CARS*. 2009;4(Suppl 1):S171–S176.
3. Schönmeier R, et al. Automated segmentation of lateral ventricles from human and primate magnetic resonance images using cognition network technology. *Magn Reson Imaging*. 2006;24(10):1377–87.
4. Schönmeier R, et al. Automatische Segmentierung des Corpus Callosum aus sagittalen Schichten von kernspintomographischen Datensätzen. In: *Proc BVM*; 2007. p. 389–93.
5. Elter M, et al. Referenzdaten für die computerassistierte Diagnose in der Mammographie. In: *Proc BVM*; 2007. p. 96–100.
6. Schäpe A, et al. Fraktal hierarchische, prozeß- und objektbasierte Bildanalyse. In: *Proc BVM*; 2003. p. 206–10.
7. Schmidt G, et al. Cognition network technology for automated holistic analysis in mammography. In: *Proc BVM*; 2007. p. 282–7.
8. Lladó X, et al. A textural approach for mass false positive reduction in mammography. *Comput Med Imaging Graph*. 2009;33:415–22.

An Object-oriented Library for Systematic Training and Comparison of Classifiers for Computer-assisted Tumor Diagnosis from MRSI Measurements

Frederik O. Kaster^{1,2}, Stephan Kassemeyer¹, Bernd Merkel³, Oliver Nix²,
Fred A. Hamprecht¹

¹Ruprecht-Karls-Universität, Heidelberg

²Deutsches Krebsforschungszentrum, Heidelberg

³Fraunhofer MEVIS, Institut für Bildgestützte Medizin, Bremen

`frederik.kaster@iwr.uni-heidelberg.de`

Abstract. We present an object-oriented library for the systematic training, testing and benchmarking of classification algorithms for computer-assisted diagnosis tasks, with a focus on tumor probability estimation from magnetic resonance spectroscopy imaging (MRSI) measurements. In connection with a graphical user interface for data annotation, it allows clinical end users to flexibly adapt these classifiers towards changed classification tasks and to perform quality control of the results. This poses an advantage over previous classification software solutions, which had to be manually adapted by pattern recognition experts for every change in the data acquisition protocols. In this article, we concentrate on software architecture and design principles of this library.

1 Introduction

In the past years, pattern recognition methods have proved their efficacy for the automated tumor detection based on MRSI measurements [1]. They often show higher robustness against measurement artifacts and noise than tumor detection methods that depend on the explicit quantitation of relevant metabolites. However, the use of quantitation-based algorithms in the clinic is already well-established, since they are typically included in the software packages supplied by MR scanner manufacturers. Furthermore, several stand-alone software products such as LCModel [2], jMRUI [3] or MIDAS [4] exist. In contrast, the application of pattern recognition-based methods still has to gain ground in clinical routine, and existing software products such as CLARET [5] or HealthAgents [6] are more of prototypical character. We believe that this may be partially due to differences in the flexibility with which both categories of algorithms can be adjusted to different experimental conditions (e.g., changes in scanner hardware and in measurement protocols) or to a different imaged organ. For quantitation-based methods one must only update the metabolite basis spectra to a given experimental setting, which can be achieved by quantum-mechanical simulation. On

the other hand, methods based on pattern recognition achieve their robustness by foregoing an explicit data model. Instead they learn the mapping from observations to tumor class labels from training examples provided by human expert annotators, which have to be provided anew for every change in experimental conditions. Since there exist many different classification techniques whose relative and absolute performance on a given task cannot be predicted beforehand, for every change in conditions a benchmarking experiment as in [1] should also be conducted to select the best classifier and monitor the classification quality.

While we cannot obviate the need for classifier retraining, benchmarking and quality assessment, we have designed an object-oriented C++ library which assists this task better than existing software. Extensibility was an important design criterion: by providing abstract interfaces for classifiers, data preprocessing procedures and evaluation statistics, user-defined classes may be plugged in with moderate effort. Hereby it follows similar ideas as general purpose classification frameworks such as Weka (<http://www.cs.waikato.ac.nz/ml/weka/>), TunedIT (<http://tunedit.org/>) or RapidMiner (<http://www.rapid-i.com>). However, it is much more focused in scope and tailored towards medical diagnostic applications. In contrast to an earlier conference contribution [7] which focused on the validation results on prostate MRSI measurements acquired at 1.5 Tesla, this paper outlines the software design architecture.

2 Materials and Methods

The library is designed for the following use case: the user defines several classifiers to be compared, the free classifier-specific parameters to be adjusted in parameter optimization and custom preprocessing steps for the data. A training and test suite is then defined, which may contain several classification tasks, e.g. predicting the voxel class (tumor vs. healthy) and the signal quality (good vs. poor). Every classifier is assigned to a preprocessing pipeline, which transforms the observations into training and test features. Some elements of this pipeline may be shared across several classifiers, while others may be specific for one classifier. Input data (observations and labels) are passed, preprocessed and partitioned into cross validation folds, and the parameters of every classifier are optimized on one fold by maximizing an estimate for the generalization error. The mean and variance of all evaluation statistics are then estimated using cross-validation, and statistical tests are conducted to decide whether the classifiers differ significantly in performance. Typically not only two, but multiple classifiers are compared against each other, which must be considered when judging significance: e.g. for five classifiers with equal performance, we have ten pairwise comparisons and a significant difference ($p_{\text{raw}} < 0.05$) is expected to occur with a probability of $1 - 0.95^{10} \approx 0.40$. Hence the “raw” p -values from statistical tests must be corrected accordingly. Finally the classifiers are retrained on the total data for predicting the class of unlabeled examples. Trained classifiers and test results may then be saved and reloaded for persistence.

3 Results

We developed a C++ library for managing multiple classifiers, preprocessing procedures and evaluation statistics: abstract base classes for all these entities provide flexibility for later adjustments. The class architecture is depicted in Fig. 1. Concerning the classifiers, we implemented bindings for support vector machines, random forests, ridge regression and principal components regression. External libraries such as LIBSVM (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>) or VIGRA (<http://hci.iwr.uni-heidelberg.de/vigra/>) are linked to provide the underlying classification algorithms. As exemplary preprocessing and feature extraction steps, we implemented a MRSI specific preprocessor that performs e.g. the extraction of relevant parts from the spectrum and that is contained in all pipelines, a whitening transformation mainly for use with SVMs and a singular value decomposition transformation for use with regression-based classifiers. Possible evaluation statistics are e.g. precision, recall and ROC curves. For p -value-adjustment, we provide McNemar's test and a recently proposed adjusted t -test [8] with multiple comparison adjustment of the p -values by Holm's step-down or Hochberg's step-up method as in [9].

All classifiers, all preprocessors and all statistics pertaining to a certain classifier are controlled by a dedicated manager object, which enforces consistency. Furthermore, every classifier encapsulates a parameter manager object controlling which parameter combinations are tested during parameter optimization.

Data input and output for training, testing and storing of the classifiers can be customized by passing user-defined input and output function objects; we use HDF5 as main storage format. For integration into a user interface, other function objects may be passed that can report progress or status information or listen for abort requests at regular checkpoints.

In order to further aid the clinical users in spectrum annotation, we developed a graphical user interface in MeVisLab (<http://www.mevislab.de>) that displays MRSI spectra from a selected slice in the context of its neighbor spectra, which can then be labeled on an ordinal scale and imported into the classification library. Hence the classifier can be flexibly adapted to new experimental protocols: Clinical end users only need to load the training data sets, manually annotate the spectra in the GUI, save the annotations and start a training and testing suite. Since in practice many spectra are not evaluable owing to excessive noise or the presence of artifacts, we enable the users to define two independent labels for each spectrum, which encode voxel class and signal quality. By training two independent classifiers for both tasks (as in [5]), it is afterwards possible to judge the reliability of the automated classification of new examples.

The library was validated with 1.5 Tesla MRSI data of prostate carcinomas (both for voxel class and signal quality classification). The outcome was already discussed in [7]: for signal quality prediction with 36864 training examples, correct classification rates (CCR) of 96.5 % – 97.3 % and area under the ROC curve values of 98.9 % – 99.3 % could be achieved by the different classifiers; and for voxel class prediction with 2746 training examples, CCRs of 90.9 % – 93.7 % and AUCs of 95 % – 98 % were obtained.

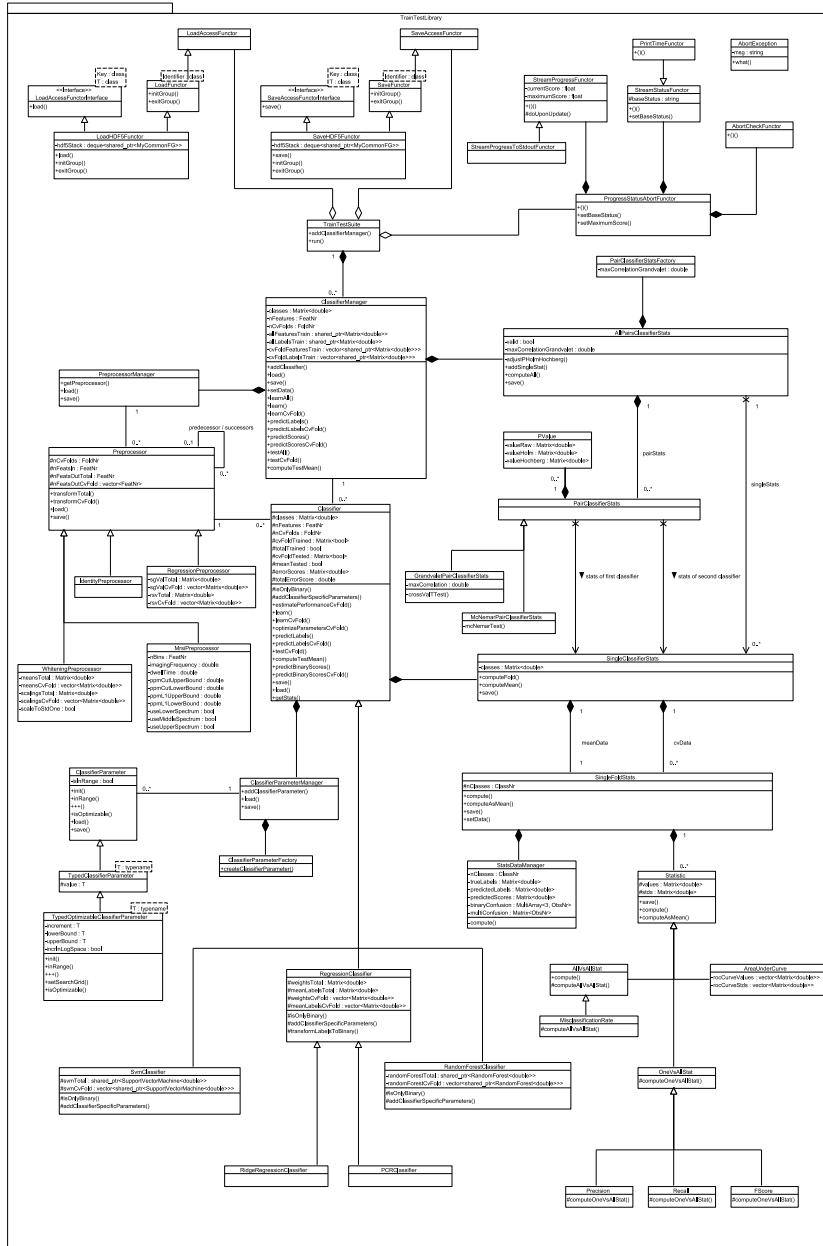


Fig. 1. Simplified UML class diagram of our library. Note that it comprises functionality for classification and managing the associated parameters (lower left), preprocessing and feature extraction (left), data input and output (top), and evaluation statistics (right and bottom right). The functors for loading and saving data, and streaming progress and status information (top part) allow the integration into user interfaces. We recommend to view this figure in the electronic version of the article.

4 Discussion

To our best knowledge, this is the first C++ library specifically designed for medical applications which allows principled comparison of classifier performance and significance testing. We believe that this will help automated quality assessment and the conduction of clinical studies. This statistical characterization functionality sets our software apart from e.g. HealthAgents [6] or the system in [10].

The adaptation to 3 Tesla MRSI measurements of brain and prostate carcinomas is ongoing. Furthermore this software will eventually be integrated into the RONDO software platform for integrated tumor diagnostic and radiotherapy planning (<http://www.projekt-dot-mobi.de>). Stringent quality assessment will be required for clinical application: the software has to be tested and reviewed extensively, and the correctness of the spectral annotations must be confirmed via histopathological assessments by several independent pathologists as in [1].

Acknowledgement. We acknowledge fruitful discussions with Ralf Floca and Ullrich Koethe. The graphical user interface was initiated by Markus Harz. This research was supported by the Helmholtz International Graduate School for Cancer Research and by the BMBF within the DOT-MOBI project (grant no. 01IB08002).

References

1. García-Gomez J, Luts J, Julià-Sapé M, et al. Multiproject-multicenter evaluation of automatic brain tumor classification by magnetic resonance spectroscopy. *Magn Reson Mater Phys.* 2009;22:5–18.
2. Provencher S. Automatic quantitation of localized *in vivo* ^1H spectra with LCModel. *NMR Biomed.* 2001;14(4):260–4.
3. Stefan D, Cesare FD, Andrasescu A, et al. Quantitation of magnetic resonance spectroscopy signals: the jMRUI software package. *Meas Sci Technol.* 2009;20:104035.
4. Maudsley A, Darkazanli A, Alger J, et al. Comprehensive processing, display and analysis for *in vivo* MR spectroscopic imaging. *NMR Biomed.* 2006;19:492–503.
5. Kelm B, Menze B, Neff T, et al. CLARET: a tool for fully automated evaluation of MRSI with pattern recognition methods. *Proc BVM.* 2006; p. 51–5.
6. González-Vélez H, Mier M, Julià-Sapé M, et al. Health agents: distributed multi-agent brain tumor diagnosis and prognosis. *Appl Intell.* 2009;30:191–202.
7. Kaster F, Kelm B, Zechmann C, et al. Classification of spectroscopic images in the DIROlab environment. In: *Proc IFMBE.* vol. 25/V; 2009. p. 252–5.
8. Grandvalet Y, Bengio Y. Hypothesis testing for cross-validation. *Département d'Informatique et Recherche Opérationnelle, University of Montréal*; 2006. TR 1285.
9. Demšar J. Statistical comparisons of classifiers over multiple data sets. *J Mach Learn Res.* 2006 Dec;7:1–30.
10. Neuter BD, Luts J, Vanhamme L, et al. Java-based framework for processing and displaying short-echo-time magnetic resonance spectroscopy signals. *Comput Methods Programs Biomed.* 2007;85:129–37.

GPU-accelerated Rendering for Medical Augmented Reality in Minimally-invasive Procedures

Matthias Wiczorek¹, André Aichert¹, Oliver Kutter¹, Christoph Bichlmeier¹,
Jürgen Landes², Sandro Michael Heining², Ekkehard Euler², Nassir Navab¹

¹Chair for Computer Aided Medical Procedures (CAMP)
Technische Universität München, Germany

²Ludwig-Maximilians-Universität München, Chirurgische Klinik und Poliklinik
Innenstadt AG CAS, Germany
wiczore@cs.tum.edu

Abstract. Recent advances in GPU programmability and performance have enabled development of real-time high quality volume visualization algorithms. Medical augmented reality systems can benefit from these developments. Task-specific visualization aids physicians in better understanding the patient’s anatomy and supports navigation of medical instruments in absence of a direct line of sight in minimally-invasive procedures. In this paper we present our results of integration of a hardware accelerated volume renderer into a medical augmented reality framework using a video see-through head mounted display (HMD). The performance of the system is evaluated in an experiment for two human CT datasets. Compared to the literature, our approach allows direct real-time stereo visualization of volumetric medical data on a HMD without prior time consuming pre-processing or segmentation. To further improve the visual perception and interaction of real and virtual objects, the renderer implements a virtual mirror and occlusion handling with the physicians hands and tracked medical instruments.

1 Introduction

Augmented reality (AR) was introduced as an alternative to monitor based visualization for the presentation of medical image data during surgery. The key task of any augmented reality system is to provide its user with good perception of the scene, especially the correct visual cues at the right time. Existing solutions for medical augmented reality often avoid direct volume rendering (DVR) and fall back to simple rendering techniques, e.g. wireframe or single slice display, or reduced rendering quality in order to maintain real-time performance. In addition many methods require time consuming pre-processing of the volume data. More advanced approaches enhance 3D perception within their stereo systems i.e. viewing windows [1] into the body. In the recent years, GPU-accelerated ray-casting has emerged as the de-facto standard algorithm for DVR [2]. ClearView by Krüger et al. improves understanding and insight of the 3D data data by

extracting focus and context (F+C) layers on the fly and clever compositing [3]. In [4] the potential of F+C rendering for medical AR is presented. F+C rendering is employed to intelligently embed the virtual scene parts in the video image within the AR scene. In [5] virtual mirror (VM) rendering for in-situ AR is proposed to overcome perspective limitations in navigated surgery settings.

This work summarizes efficient implementations for GPU-accelerated direct volume rendering (DVR) and proposes smart extensions to get the best visual perception while ensuring the required real-time update rate. Important visual cues, such as shape from shading, depth from occlusion or motion parallax can greatly improve user perception of the AR scene.

2 Methods and Materials

In this work we extend and improve the volume rendering and F+C techniques within the AR scene [6], by additional techniques in order to support several new in-situ visualization features. The AR system for optical tracking and the video see-through HMD for in-situ visualization was originally developed by Sauer et al. [7]. Tracking of the objects in the scene is accomplished by two separate optical tracking systems. An outside-in optical tracking system consisting of four infrared ARTtrack2 (A.R.T. GmbH, <http://www.ar-tracking.de/>) cameras mounted to the ceiling of the room and an inside-out optical tracking system using an IR camera mounted directly on the HMD. Position and orientation of the users head are accurately determined by the inside-out tracking system, while the outside in tracking covers a much greater tracking volume and is used for all tracked objects in the scene. A reference system composed of a known arrangement of optical tracking markers, visible by both tracking systems serves as the common reference frame of the scene. For more details on the used system setup, and calibration of tracking targets (e.g., phantom, instrument), see [6].

In contrast to [6] we rearranged the rendering pipeline. In the first stage a first hit iso surface ray-casting extracts a certain surface (e.g. skin or bone). This hit texture should be extracted before all other rendering passes, because for several rendering techniques this hit texture has to be fully defined and thus not be effected by occlusion handling. In the second stage of the pipeline the hand occlusion is implemented. In the refined version the size of the focus region and hit texture positions are incorporated to improve detection stability. The following stages implement volume rendering and additional effects, e.g. virtual mirror, MPRs (multi-planar reformations). As proposed by Fischer et al. [8] the hit texture is used for occlusion handling for medical instruments. This occlusion handling is again implemented using the OpenGL shading language (GLSL). While the instrument is rendered, the parts of the instrument which are located above the skin, are replaced by the corresponding values read from the video camera image, while the inner parts are rendered traditionally as wireframe models. The individual results of the different rendering stages are illustrated in Fig. 1(b-d).

In order to integrate the paradigm of the virtual mirror [5] into our setup, we combined the virtual mirror and the presented DVR techniques. The mirrored image for the virtual mirror is computed by an additional raycast pass. The computational costs for this additional render pass are kept to a minimum by using the mirror as a stencil for this pass. The usability of the mirror is improved by several features like attachment to an instrument, automatic alignment to the instruments tip, different focus point for the mirror in order to provide the best visual information and linear interpolation based zooming.

Because of the popularity of MPRs amongst medical staff, we integrated additional MPR planes in the visualization of our AR system. We decided to render only the parts of the MPRs which are below the skin, thus the MPR perfectly fits into the F+C visualization. Like the virtual mirror, MPRs can be attached to an instrument as well in order to move and position the MPR within the body.

3 Experiments and Results

We evaluated the performance in an experiment for two human CT datasets, head (296x320x420) and thorax (256x368x522), extracted from the original Visible Korean Human CT data. Bichlmeier et al. [9] created a phantom based on the VKH data to facilitate repeatable realistic experiments. The life-size phantom ranges from the head to the lower hips. Exemplary result images for the head

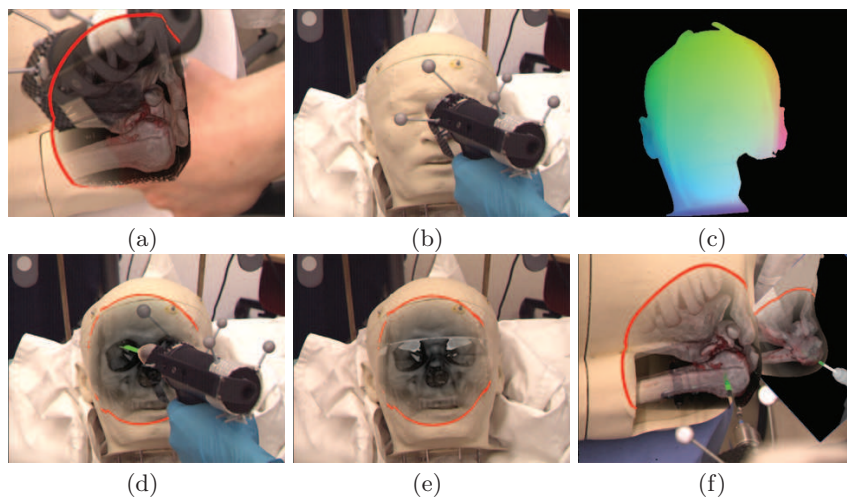


Fig. 1. (a) Illustration of the occlusion problem. (b,c,d) Render pipeline for correct occlusion handling, (b) video texture, (c) hit texture for the skin, (d) final composition of (b) and (c). Video image is used as context layer, Focus layer is rendered with volume rendering. Occlusion handling is shown for instruments and hands. (e) like (d) with in-body MPR. (f) Focus and Context rendering with shaded volume rendering for the focus layer (bone), virtual mirror and instrument.

Table 1. Average rendering performance (fps) results for visible korean head and thorax datasets over 2 minutes. (Pre-Int stands for preintegrated transferfunction) (a) Conventional DVR (b,c) F+C rendering. A video isosurface context layer is always rendered and paired with the same DVR modes as in the conventional DVR evaluation. (c) One virtual mirror and one surgical drill model added to (b).

Rendermode	(a) DVR		(b) F+C		(c) F+C, VM	
	Head	Thorax	Head	Thorax	Head	Thorax
DVR	>30	18	> 30	26	> 30	25
DVR Shaded	>30	8	> 30	12	> 30	12
Pre-Int	27	18	> 30	25	> 30	25
Pre-Int Shaded	13	7	28	16	28	15

and thorax datasets of the visible korean human phantom are shown in Fig. 1(d-f). For the performance measurements the internal render target resolution was set to the camera resolution (640×480), the viewport resolution was set to the HMD resolution (1024×768). Empty space leaping and early ray termination optimizations were activated. A PC workstation with a NvidiaTM Geforce GTX 275 is used to run the renderer. For the evaluation a volunteer wore the HMD and inspected the phantom by continuous motion around it for two minutes. The average framerate was measured using Fraps (<http://www.fraps.com/>), a freely available benchmarking tool. Note that the maximum framerate of the system is limited by the update rate of the slowest component in the system. Thus, the average framerate can never be more than 30 frames per second, the analog video camera refresh rate, in our experiments.

The measured average framerate for various DVR rendering modes is depicted in Tab. 1(a), for F+C rendering in Table 1(b) and for F+C rendering with virtual mirror and instrument in Tab. 1(c). The better performance of the F+C rendering modes compared to the DVR rendering modes can be explained by the reduced number of pixels for which rays are cast through the dataset when the focus region optimization is enabled.

4 Discussion

This paper discusses the integration and possibilities of GPU-accelerated DVR in a medical augmented reality environment combined with support for virtual mirror, occlusion handling for physician hands and medical instruments and in-body augmentation of medical instruments. The demonstrated techniques are an important step towards the integration of in-situ AR systems to medical navigation applications. F+C volume visualization not only provides improved visual perception, but also aids in maintaining real-time performance, compared to conventional DVR, as demonstrated by the experiments.

Ongoing work is the evaluation of the system in simulated surgical navigated procedures together with our clinical partners. Future work will focus on (i)

technical improvement of the system, e.g. real-time depth map reconstruction using the stereo camera setup, and (ii) automatic medical workflow optimized visualization mode selection. Therefore, the specific requirements have to be analyzed for each procedure over the complete workflow, from pre-operative imaging till the end of the intervention. For these dataset the visualization modes will be iteratively refined based on feedback from our medical partners.

References

1. Bajura M, Fuchs H, Ohbuchi R. Merging virtual objects with the real world: seeing ultrasound imagery within the patient. In: Proc SIGGRAPH; 1992. p. 203–10.
2. Engel K, Hadwiger M, Kniss JM, et al. Real-Time volume graphics. AK Peters, Ltd.; 2006.
3. Krüger J, Schneider J, Westermann R. ClearView: an interactive context preserving hotspot visualization technique. *IEEE Trans Vis Comput Graph*. 2006;12(5).
4. Bichlmeier C, Wimmer F, Heining SM, et al. Contextual anatomic mimesis: hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality. In: Proc ISMAR; 2007. p. 129–38.
5. Bichlmeier C, Sielhorst T, Navab N. The tangible virtual mirror: new visualization paradigm for navigated surgery. In: Proc AMI-ARCS Workshop; 2006.
6. Kutter O, Aichert A, Bichlmeier C, et al. Real-time volume rendering for high quality visualization in augmented reality. In: Proc AMI-ARCS Workshop; 2008.
7. Sauer F, Khamene A, Bascle B, et al. A head-mounted display system for augmented reality image guidance: Towards clinical evaluation for iMRI-guided neurosurgery. In: Proc MICCAI; 2001. p. 707–16.
8. Fischer J, Bartz D, Straßer W. Occlusion handling for medical augmented reality using a volumetric phantom model. In: Proc VRST; 2004. p. 174–7.
9. Bichlmeier C, Ockert B, Kutter O, et al. The visible Korean human phantom: realistic test & development environments for medical augmented reality. In: Proc AMI-ARCS Workshop; 2008.

SASOMI

Semi-Automatic Segmentation of Medical Images

Michael Emmersberger¹, Stefanie Demirci¹, Reza Ghotbi², Nassir Navab¹

¹Computer Aided Medical Procedures (CAMP), TU München

²Vascular Surgery Department, Kreisklinik München-Pasing

emmersbe@in.tum.de

Abstract. This paper presents an interactive graphical user interface that provides an intuitive solution for refining an automatic pre-segmentations of abdominal aortic aneurysms. With current applications the segmentation of aneurysm thrombus has to be done completely manual without the possibility to refine automatic pre-segmentations. The SASOMI software therefore provides interpolation methods based on NURBS surfaces combined with edge fitting to accelerate the process of segmentation. The implementation was done using C++, Qt and OpenGL. Robustness and user friendliness have been tested in a set of different experiments.

1 Introduction

Many clinical applications involve exact segmentation of anatomical structures in all kind of 3D images. When automatic methods fail to meet the required robustness, user interfaces that allow for refining automatic segmentation results are of great help for the medical staff.

In the current paper, we focus on abdominal aortic aneurysm (AAA). It is a permanent, irreversible, localized dilatation in the abdominal section of the aorta. In most cases, an aneurysm thrombus is present caused by a rupture of the most inner layer of the aortic wall through. This allows blood to enter in between the two layers and coagulate.

For diagnosis and detection of AAA, computed tomography angiography (CTA) is the standardized imaging modality. The aortic lumen is highlighted by injected contrast agent and has strong gradient values to surrounding structures. The aortic wall and aneurysm thrombus, however, are visually very hard to detect without any anatomical expertise. Additional difficulties for segmentation algorithms are introduced by severe calcification occurrence inside the thrombus.

For allowing an optimal treatment planning of AAAs, the focus of medical as well as engineering research goes towards integrating biomechanical models in the medical workflow. In order to give realistic simulation results, the geometry of the aneurysm and its thrombus need to be extracted as accurate as possible. In the last decade, several approaches have been published allowing fully- or semi-automatic AAA segmentation [1, 2]. Despite of their promising results,

it has been shown that in clinical practice, these methods fail to provide the required robustness.

Graphical user interfaces for segmentation of anatomical structures have become extremely popular in clinical practice. Here, a variety of commercial and non-commercial software is available for general usage [3, 4] as well as for specific medical applications [5]. In the case of abdominal aortic aneurysms a CTA scan can consist of up to 1000 slices. Therefore a segmentation tool needs to support some kind of interpolation in-between the slices as well as post-refinement possibilities, to improve usability. Furthermore intensity region growing is not applicable as a semi-automatic approach, since the ROIs are not always surrounded by clear and closed contours. To the best of our knowledge, there is no such refinement software available for the segmentation of AAAs.

In this paper, we present an interactive graphical user interface that provides for the medical staff an intuitive solution for refining an automatic AAA pre-segmentation. The software transforms the pre-segmented volume in a NURBS surface providing an intuitive way of interacting with already present segmentation.

2 Materials and Methods

As displayed in Fig. 1, the software consists of four different views: axial, sagittal, coronal and volume 3D. While each of the first three gives a slice based view of the loaded image according to its direction, the Volume view shows a 3D visualization of the segmentation that can be zoomed and rotated. The segmentation tree underneath the main view offers possibilities to change color and opacity of all

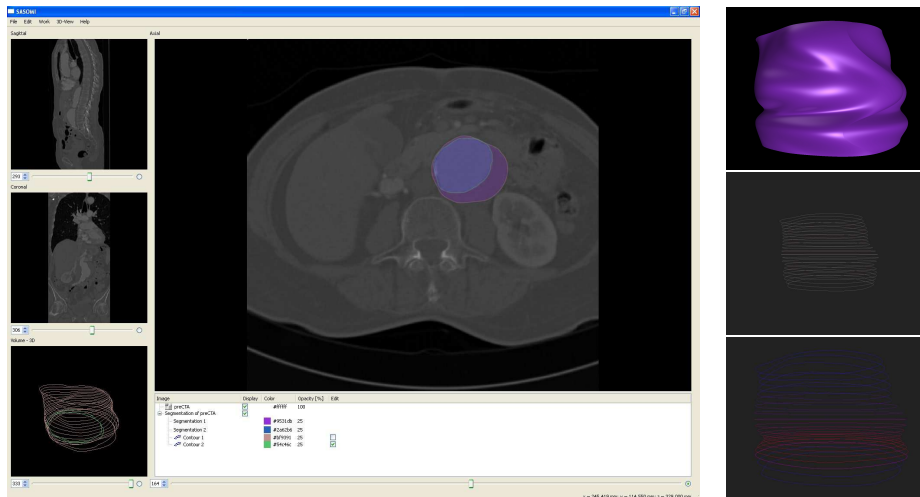


Fig. 1. SASOMI software. Left: overall layout; right: three volume views.

segmentations. In the following, we describe some of SASOMI's main features that also represent the main contribution of this paper

- *Tracing*: In addition to free-hand-drawing of a contour in a 2D slice, SASOMI supports edge tracing that simplifies following the outline of an edge using edge detection. While in tracing mode, edge points that are detected in the nearby area of the mouse pointer are displayed as yellow dots (Fig. 2). To refine an already present contour it is possible to start tracing at one point of the contour and finish at second point. The part in between will then be substituted by the new one. Erasing the last drawn points can be done by drawing backwards. Thanks to edge detection it is not necessary to hit these points exactly.
- *Slice interpolation*: Since it is quite time-consuming to do segmentations on every slice, SASOMI computes a NURBS surface [6] from already segmented slices which is then used to generate the missing slices in-between. It also closes gaps in segments that lie within the same slice. Already segmented slices are transformed into a NURBS curve for a more accurate and smoother visualization.
- *Edge fitting*: To refine the interpolation result this function fits the segmented points to the closest edge found in the image. The calculation of the edge points itself uses a quadratic 2D facet model on the gradient image. Compared to region growing one could consider the contour resulting from the interpolation as a seed which is in most cases already quite close to the actual gradient in the image. Accordingly a lot less comparisons are needed and the given shape also helps the algorithm to stay in a certain boundary.
- *Directional edge fitting*: A more enhanced version of the popular edge fitting method. Edge points are captured only in the approximate direction of the curve's normal (Fig. 2, left). In addition, a search along the new edge is performed and while points that are not likely to be part of this edge are removed, missing edge points are included. The result will then be transformed into a NURBS curve (Fig. 2, right). Compared to the standard Edge Fitting method it can return more accurate results especially in corners.

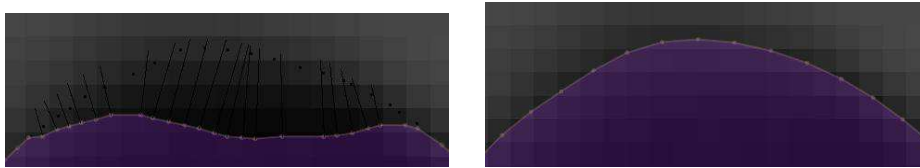


Fig. 2. Edge fitting.

Table 1. Evaluation (R=Robustness, U=Usage, IT=Interaction tools, TE=Time effort, V=Visualisation).

	Task1					Task2				
	R	U	IT	TE	V	R	U	IT	TE	V
E1	10	10	9	6	5	10	10	9	6	5
E2	10	9	8	5	5	10	10	10	5	5
E3	10	9	8	6	7	10	7	9	6	7
M1	10	7	9	10	5	10	7	10	10	5
M2	10	9	10	10	6	10	9	10	10	6
M3	10	9	9	9	6	10	9	8	9	6

3 Results

The SASOMI software was implemented in C++ using a Qt-based User Interface and OpenGL as graphics engine. We tested its robustness and user friendliness in a set of different experiments. We asked three engineers and three medical doctors to evaluate the software in two steps. The first task consisted of loading an abdominal CTA scan of a patient suffering from AAA and segmenting the AAA thrombus manually with the help of the given interaction tools. Secondly, these presegmentations were mixed and each user was given one segmented volume created by a different user. They were then asked to use our software to refine this segmentations.

The evaluation results are presented in Tab. 1. The users had to rank the software in each category in numbers 1 (bad) to 10 (excellent).

While the other categories were found to be good to excellent, all users complained about the insufficient volume views. All medical doctors stated that such a volume view of the segmentation does only make sense within a volume rendering environment of the surrounding anatomical structure. The engineers criticized the missing mesh visualization that would make more sense for accessing the geometry in detail.

To test the improvement of interaction time, gained by using the implemented functions, we selected several persons with different level of experience considering the use of graphics software or SASOMI itself. While beginners are users with little or no experience, three users in our test are working as graphic- or web designers. Both got a short introduction in the program's user interface. The persons called experts are using SASOMI already for several weeks.

The task was to segment an object over 20 slices using the different functions the program provides. It is quite obvious that edge detection already accelerates the segmentation process. In addition the variance of the results between the different users is reduced. With interpolation it was enough to process only every fifth slice. The interpolation time of 0.7 seconds for this scenario is already included in the results. It should also to be mentioned that even with nurbs

Table 2. Interaction time.

User Level	Beginner			Experienced			Expert		
User	1	2	3	4	5	6	7	8	9
Free Hand	6:22	4:52	5:12	3:56	4:13	4:28	4:09	3:58	3:26
Edge Detection	2:48	3:02	3:48	2:26	1:59	2:08	2:09	1:57	1:49
Edge Detection +Interpolation	0:50	0:57	1:02	0:52	0:46	0:42	0:46	0:35	0:34

surfaces interpolation does not produce perfect results for every object. Edge fitting can then be used to reduce the effort of refinement.

4 Discussion

The advantages of SASOMI clearly lie in facilitating the process of segmentation and providing an intuitive way of interacting with already present segmentation. Additionally the 3D-model enables a more realistic impression of the segmented areas. Although SASOMI has been developed in the context of abdominal aortic aneurysms it is applicable for a variety of fields. For example it is currently used for segmentation of midbrain area in ultrasound volumes.

As future work, we plan to integrate the user feedback and adapt the volume visualization to be more usable. Also, a further adaptation of the software and an optimization of the internal algorithms for ultrasound images is already planned. Another idea is integrating new and alternative interaction methods like haptic devices or 3D displays. Providing an intuitive set of adjustable parameters could make the edge fitting function adaptable for a larger variety of situations. So far the application has only been tested with Windows XP, but portability is given through the use of C++, Qt and OpenGL for implementation.

References

1. Demirci S, Lejeune G, Navab N. Hybrid deformable model for aneurysm segmentation. In: Proc IEEE ISBI; 2009. p. 33–6.
2. Olabbarriaga SD, Rouet JM, Fradkin M, et al. Segmentation of thrombus in abdominal aortic aneurysms from CTA with nonparametric statistical grey level appearance modeling. *IEEE Trans Med Imag.* 2005;24(4):477–85.
3. Wolf I, Vetter M, Wegner I, et al. The medical imaging interaction toolkit. *Med Image Anal.* 2005;9(6):594–604.
4. Mimics [program on internet]; 2009. Available from: www.materialise.com/mimics.
5. König L, Groher M, Keil A, et al. Semi-Automatic segmentation of the patellar cartilage in MRI. In: Proc BVM; 2007. p. 404–8.
6. Piegl L, Tiller W. *The NURBS Book*. Springer; 1997.

Ein System zur berührungslosen, volumetrischen Vermessung von Gesichtsschwellungen

Christoph John, Ulrich Schwanecke

Fachbereich für Design Informatik Medien, Hochschule RheinMain
Christoph_John@gmx.de

Kurzfassung. Dieser Beitrag beschreibt ein kostengünstiges Verfahren zur berührungslosen Vermessung von Gesichtsschwellungen im zahnmedizinischen Umfeld. Zielsetzung des entwickelten Systems ist die Vermessung und Dokumentation von Therapie und Heilungsprozessen nach operativen Eingriffen. Die vorgestellte Anwendung basiert auf einem handelsüblichen Streifenlichtscanner zur Akquisition von Tiefendaten, aus denen eine volumetrische Rekonstruktion des Patientengesichts erzeugt wird. Zur Dokumentation von Heilungsprozessen werden Patienten über einen postoperativen Zeitraum hinweg mehrfach gescannt. Die angefertigten Scans werden anschließend paarweise registriert und Differenzvolumina von Schwellungsregionen semiautomatisch vermessen.

1 Einleitung

Im Rahmen der Evaluierung und Analyse von Therapiemethoden, unter anderem im zahnmedizinischen Umfeld, besteht die Notwendigkeit zur detaillierten Vermessung und Dokumentation von Gesichtsschwellungen.

Schon seit längerem existieren Systeme zur exakten optische Aufzeichnung und Vermessung von Körperteilen, wie beispielsweise der Body-Scanner [1] oder der Bodypart-Scanner [2]. Die Einsatzgebiete dieser Messsysteme liegen bisher jedoch nahezu ausschließlich in Bereichen der Computergrafik und Animation, sowie in jüngster Zeit auch in der Bekleidungsindustrie. Gründe dafür finden sich in den oft hohen Anschaffungskosten, dem für die Installation notwendigen Platzbedarf und in der mangelnden Anpassung an die Bedürfnisse im medizinischen Umfeld. Dies hat zur Folge, dass auch heute noch medizinische Studien mit Schublehre und Maßband durchgeführt werden.

Die hier vorgestellte Software ermöglicht die berührungslose, volumetrische Vermessung und Dokumentation von Gesichtsschwellungen, wie sie etwa nach zahnmedizinischen Eingriffen auftreten. Dabei ist sie speziell auf die Bedürfnisse im medizinischen Umfeld zugeschnitten und basiert ausschließlich auf kostengünstigen Standardkomponenten.

Im folgenden Kapitel werden wir den Systemaufbau unserer Anwendung sowie ihre Einzelkomponenten vorstellen. Dem folgt eine Diskussion erster Messergebnisse, und ein Ausblick auf geplante Weiterentwicklungen.

2 Das System

Die vorgestellte Anwendung verwendet einen kalibrierungsfreien Streifenlicht-scanner der aus einem handelsüblichen Videoprojektor und einer monochromen USB Kamera besteht. Beide sind im Einzelhandel für unter 2000 Euro erhältlich. Der Scanner erfasst ein Arbeitsvolumen von etwa 40 cm^3 bei einer Distanz von 110 cm. Eine Installation ist daher an nahezu jedem Büroarbeitsplatz ohne nennenswerten Platzbedarf und Zusatzkosten möglich.

Die Software des vorgestellten Systems basiert auf einer MySQL Datenbankkomponente sowie einem Scan-Editor und einer Scanner-, Vermessungs- und Messdatenansicht. Die Datenbankkomponente (Abb. 1(a)), verwaltet dabei alle Patienteninformationen, Anamnesedaten sowie Registrierungs- und Messdaten und ermöglicht damit einen standardisierten Zugriff auf alle Anwendungsdaten auch für externe Software. Die verbleibenden Komponenten ermöglichen die Durchführung einer Schwellungsmessung und werden im Folgenden anhand eines typischen Messverlaufs vorgestellt.

Die Schwellungsvermessung beginnt mit dem Scannen des Patienten. Hierzu benötigt der Scanner, der direkt aus der Anwendung heraus angesteuert wird, etwa eine Sekunde (Abb. 1(b)). Aus den aufgezeichneten Daten heraus stellt der Anwender zunächst die Gesichtsregion des Patienten frei. Dazu steht ihm eine Editorkomponente zur Verfügung, die einfache Werkzeuge zur Selektion und zum Löschen von selektierten Daten anbietet (Abb. 2(a)). Die verbleibenden Daten werden von der Anwendung zur Generierung eines polygonales Netzes verwendet, welches nun gegen andere Scans registriert werden kann.

Die Registrierung von Scans wird in zwei Schritten vollzogen, der Vor- und Feinregistrierung. Für die Vorregistrierung stellt die Anwendung ein automatisches und ein manuelles Verfahren zur Verfügung. Das automatische Verfahren ist häufig ausreichend und basiert auf der Schwerpunktüberlagerung der zu registrierenden Scans. Werden Scans mit geringem Überlappungsbereich gegeneinander registriert, muss jedoch auf eine manuelle Ausrichtung zurückgegriffen

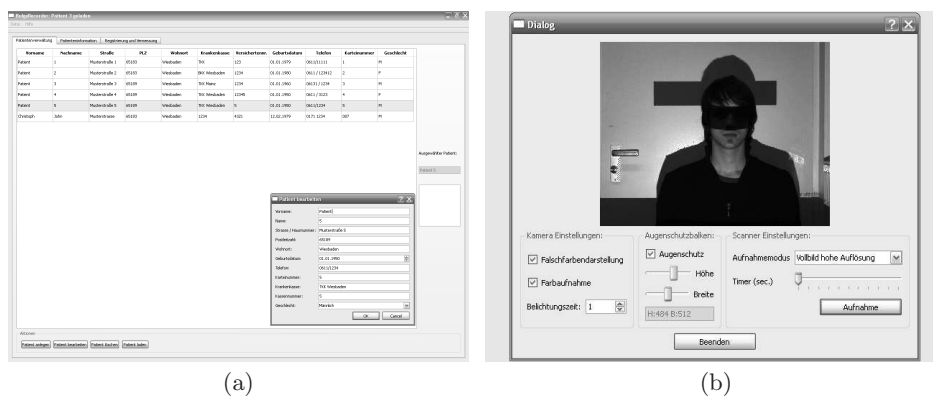


Abb. 1. (a): Datenbankansicht der Anwendung. (b): Ansicht des Scannerdialogs.

werden. Zur manuellen Ausrichtung gibt der Anwender vier Punktkorrespondenzen zwischen den Scans vor, die dann überlagert werden (Abb. 2(b)). Die grob ausgerichteten Scans können nun mit einem „Iterative Closest Points“ (ICP) Verfahren [3, 4] genauer gegeneinander registriert werden. Dazu werden wiederholt nächste Nachbarn zwischen Scans als Punktkorrespondenzen selektiert, um diese mit pro Iteration steigender Genauigkeit gegeneinander zu registrieren.

Zwischen den registrierten Scans können nun Schwellungsvolumina semiautomatisch vermessen werden. Dazu gibt der Anwender per Mausklick einen Saatpunkt für eine „Region Growing“ Prozedur [5] vor. Diese findet automatisch Schwellungsgrenzen in den Scans und startet eine Diskretisierungsprozedur auf dem Differenzvolumen mit vorher konfigurierbarer Präzision (Abb. 3). Die hier skizzierte Vermessung erfolgt also nach initialer Vorgabe einer Suchregion voll automatisch. Die manuelle Initialisierung erlaubt dabei dem Anwender die Zu-

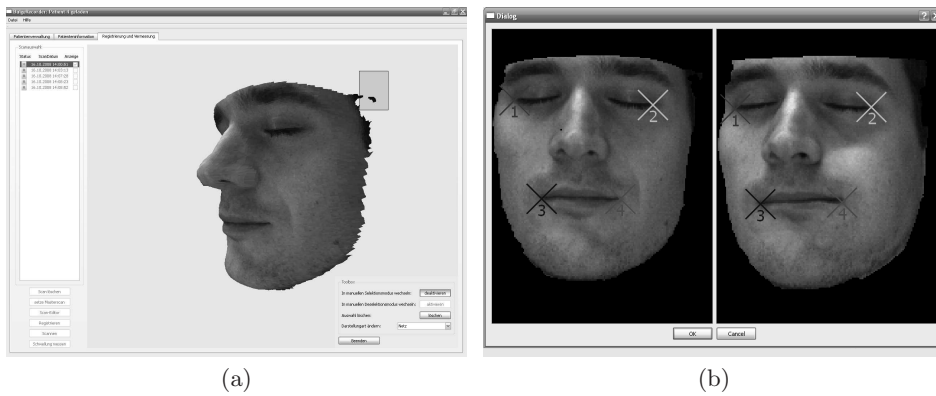


Abb. 2. (a): Ansicht der Editorkomponente. (b): Ansicht des manuellen Registrierungsdialogs.

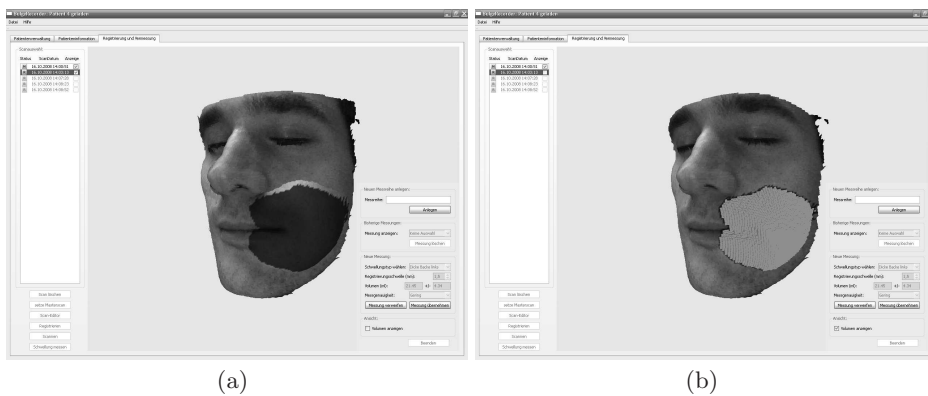


Abb. 3. (a): Ein Scan mit markierten Schwellungsbereich. (b): Ein Scan überlagert mit diskretisiertem Schwellungsvolumen.

ordnung der Messung zu einer definierten Messreihe. Damit wird die parallele Verwaltung von mehreren Messreihen für nicht überlappende Schwellungsvolumina ermöglicht.

Ein einzelner Gesichtsscan besteht aus etwa 25000 Messpunkten. Um diese Daten schnell verarbeiten zu können und interaktive Antwortzeiten auf eine Selektion hin zu gewährleisten, verwendet die Anwendung interne Datenrepräsentationen basierend auf KD-Bäumen [6] und Half-Edge Datenstrukturen [7], sowie einen Octree [8] zur Diskretisierung des Schwellungsvolumens. KD-Bäume werden verwendet um schnell nächste Nachbarvertices zwischen Scans zu finden. Deren Abstände dienen als Terminierungskriterium für das Auffinden von Schwellungsregionen. Half-Edge Datenstrukturen auf der anderen Seite werden zum schnellen Auffinden von Nachbarvertices innerhalb eines Scans benötigt.

3 Ergebnisse

Die Ergebnisse der Vermessungen werden in einer detaillierten tabellarischen Ansicht und zusätzlich über einen zeitlichen Verlauf hinweg in einem Diagramm in der Patientenansicht dargestellt (Abb. 4). Jede Messung wird hier als vertikaler Balken repräsentiert. Dabei beschreibt die mittlere Balkenposition auf der Ordinate das geschätzte Messvolumen und die Balkenhöhe den durch die Diskretisierung entstehenden systematischen Messfehler. Dieser beträgt bei einer Messkonfiguration mit hoher Genauigkeit in etwa fünf Prozent des Gesamtvolumens. Bei Bedarf könnte dieser Fehler durch eine feinere Diskretisierung weiter reduziert werden, was jedoch nicht notwendig ist, da andere Fehlerquellen wie etwa Registrierungsfehler und die Tatsache das ein Gesicht kein rigides Objekt darstellt die Messung zusätzlich beeinflussen. Ein Diskretisierungsfehler von fünf Prozent ist daher eine hinreichend genaue Kenngröße.

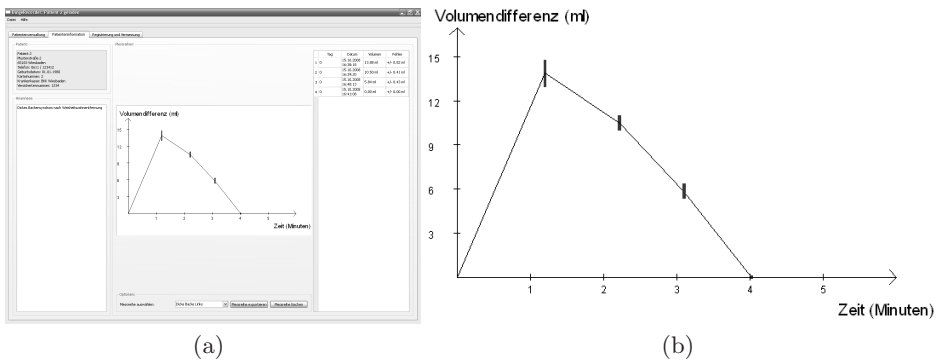


Abb. 4. Messergebnisse: (a): Zeigt Messergebnisansicht der Anwendung in der Messungen, Anamnesedaten und Patienteninformationen gemeinsam dargestellt werden. Messergebnisse können von hier zur weiteren Verarbeitung exportiert werden. (b): Vergrößerter Ausschnitt der Messergebnisansicht. Einzelmessungen werden als Balken dargestellt deren Längen den systematischen Messfehler repräsentieren.

4 Diskussion

Es wurde ein optisches Verfahren zur berührungslosen, volumetrischen Vermessung und Dokumentation von Gesichtsschwellungen für das zahnmedizinische Umfeld vorgestellt. Die bisher erzielten Messergebnisse weisen einen systematischen Messfehler von etwa fünf Prozent auf. Damit stellt diese Anwendung eine günstige Alternative zu traditionellen Evaluierungsverfahren im zahnmedizinischen Umfeld zur Verfügung, die bisher auf die manuelle Vermessung mit Schublehre und Maßband angewiesen sind.

Für die nächste Zeit planen wir detailliertere Genauigkeitstests verbunden mit einer ersten zahnmedizinischen Studie. Desweiteren ist die Integration eines vollautomatischen Registrierungsverfahrens geplant, bei der eine Vorregistrierung der Scan-Daten entfällt.

5 Danksagung

Wir bedanken uns herzlich bei Dres. Ralf Schulze und Dan Brüllmann von der Poliklinik für Zahnärztliche Chirurgie an der Johannes Gutenberg-Universität Mainz für die gute Zusammenarbeit und die umfangreichen Ratschläge aus zahnmedizinischer Perspektive. Teile dieser Arbeit wurden im Rahmen des Programmes PRO INNO II des Bundesministeriums für Wirtschaft und Technologie gefördert.

Literaturverzeichnis

1. Stein N, Minge B. VIRO 3D: fast three-dimensional full-body scanning for humans and other living objects. In: Proceedings of SPIE. vol. 3313. SPIE; 1998. p. 60.
2. Josten M, Rutschmann D, Massen R. Messbar einfach: Mobiles und wirtschaftliches 3D Body Scanning in der Medizin mit dem MagicalSkin Scanner^(TM). In: Proc BVM; 2003. p. 216–219.
3. Chen Y, Medioni G. Object modeling by registration of multiple range images. *Image Vis Comput.* 1992;10(3):145–155.
4. Besl PJ, McKay HD. A method for registration of 3-D shapes. *IEEE Trans Pattern Anal Mach Intell.* 1992;14(2):239–256.
5. Shapiro L, Stockman G. *Computer Vision.* 2001. Prentice Hall; 2001.
6. Bentley JL. Multidimensional binary search trees used for associative searching. *Commun ACM.* 1975;18(9):509–517.
7. Botsch M, Steinberg S, Bischoff S, et al. Openmesh-a generic and efficient polygon mesh data structure. In: *OpenSG Symposium.* vol. 2002; 2002.
8. Mehl M, Meagher D. Geometric Modeling Using Octree-Encoding. *Computer Graph Image Process.* 1982;19(2):129–147.

A Radiometry Tolerant Method for Direct 3D/2D Registration of Computed Tomography Data to X-ray Images Transfer Function Independent Registration

Boris Peter Selby¹, Georgios Sakas², Stefan Walter¹, Wolfgang-Dieter Groch³,
Uwe Stilla⁴

¹MedCom GmbH, Darmstadt

²Cognitive Computing and Medical Imaging, Fraunhofer IGD Darmstadt

³Fachbereich Informatik, University of Applied Sciences Darmstadt

⁴Photogrammetry and Remote Sensing, Technische Universität München

pselby@medcom-online.de

Abstract. As exact dose delivery is essential for radiological cancer treatment, image-guided radiotherapy (IGRT) methods are used to estimate corrections of the tumor alignment. This is often done through comparison of a computed tomography (CT) to the patient alignment visible in digital radiography images (DRs) acquired from within the treatment device. Digitally reconstructed radiography images (DRRs) computed from the CT are then geometrically registered to the DRs. A problem is that radiometric properties of DRs and DRRs can vary profoundly. If a registration algorithm does not use volumetric CT data directly it is unable to regard deviations of the physical image formation to the simulated image formation. We present a novel method allowing direct DR to CT registration. It is designed to be radiometry tolerant by adapting the simulated X-ray transfer function to observed DR intensities. This is done by solving an overdetermined system of equations given by the histograms along rays through the voxel matrix of the CT. Remaining errors serve as measure for the image dissimilarity, thus minimization in 6 degrees of freedom (DOF) gives the transformation between the images. Thereby higher radiometric tolerance can be achieved, as misalignments can be identified even if DR images are acquired with inappropriate radiation parameters.

1 Introduction and related Work

By registration of preoperative and intra operative data the misalignment of a radiation treatment target can be computed. Most approaches perform 2D-3D registration of CT to DR images by creating DRRs through projection of the CT and subsequent similarity maximization between DRRs and DRs [1, 2]. A problem is that radiometric properties of reconstructed DRRs often are totally different from the acquired DRs, as the physical image formation process can

hardly be modeled due to complicated statistical properties of photon-matter interactions (i.e. photoelectric absorption, Rayleigh scattering, etc.) [3]. E.g. [4] and [5] among others propose Mutual Information (MI) as a robust similarity measure for multi-modal 2D-2D registration, but in the case of DR to CT registration MI is not able to incorporate 3D information from the CT and performance depends mainly on the implementation of the CT projection process. In [6] an approach is described that adapts CT data to the used X-ray energy to increase radiometric independency. Whilst this method is suitable for some cases, it is strongly limited as only a single free parameter (the X-ray kilovoltage peak) is adapted during ray-tracing, instead of regarding possible variations for each attenuation value. In this contribution a new Transfer Function Independent Registration (TFIR) approach is presented, reducing dependency of the 2D-3D registration from radiometric properties. The image formation is modeled and Least-Squares Fit is used to adapt it by comparison of DR intensities to rays through the CT. A similarity measure is derived from remaining errors between observed and simulated intensities. Results of our approach are compared to MI based registration.

2 Materials and Methods

Our approach employs the Least-Squares method, based on a functional model of the X-ray image formation process.

2.1 The TFIR Algorithm

In Fig. 1 we give an overview over the complete 3D-2D registration algorithm. Assuming a certain geometric alignment for the patient, modeled by the CT,

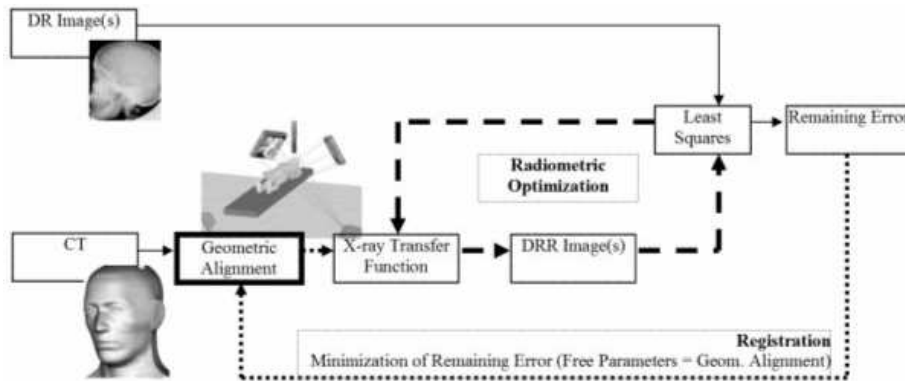


Fig. 1. The proposed 3D-2D registration algorithm, consisting of an optimization of the X-ray transfer function for image formation and the optimization of the geometric transformation in 6 DOF.

one or more DRR images are created (i.e. one per DR image), employing an X-ray transfer function, roughly emulating physical image formation. The resulting intensities are compared to the observed intensities from the DR(s). The transfer function is adapted using a constrained Least-Squares Fit. The calculation bases on an overdetermined linear system with high redundancy, and minimization of the remaining errors with the 6 parameters of rigid transformation of the patient leads to a corrected alignment.

2.2 Radiometric Optimization

First a simplified model of mono-energetic X-ray image formation based on the Lambert-Beer's Law is linearized. The photon fluence transmitted through a material is estimated by (1)

$$I_{tr} = I_0 * \exp\left(-\int_0^s \mu(\eta) d\eta\right) \quad (1)$$

where I_0 is the photon fluence that would be measured without any attenuation, s is the total length of the ray's path through matter and μ is the absorption coefficient at position η on the path. However, grey values G represent absorbed photons instead of transmitted photons in digital X-ray images so that $G = I_0 - I_{tr}$. We normalize pixel intensities G to a range between 0 and 1 and contribute to the fact that absorption values are represented in a discrete raster. Equation 2 then gives the grey value for a ray with N_s equal sized steps through the CT volume

$$G = 1 - \exp\left(-s * \sum_{i=1}^{N_s} \mu(i)\right) \quad (2)$$

Equation 3 shows linearized and simplified eq. 2 to obtain an observation b from a grey value G

$$b = \frac{-\ln(1 - G)}{s} = \sum_{i=1}^{N_s} \mu(i) \quad (3)$$

Selected sets of pixels in the X-ray images are regarded as observations. According to the linearized model they are expressed as sums over the histograms of the respective rays through the CT. The equations for the observed intensities are

$$\begin{aligned} \hat{b}_1 &= b_1 + \hat{r} = \hat{x}_1 h_{11} \mu_1 + \hat{x}_2 h_{12} \mu_2 + \cdots + \hat{x}_u h_{1u} \mu_u = f_1(\hat{x}) \\ \hat{b}_2 &= b_2 + \hat{r} = \hat{x}_1 h_{21} \mu_1 + \hat{x}_2 h_{22} \mu_2 + \cdots + \hat{x}_u h_{2u} \mu_u = f_2(\hat{x}) \\ &\vdots \\ \hat{b}_n &= b_n + \hat{r} = \hat{x}_1 h_{n1} \mu_1 + \hat{x}_2 h_{n2} \mu_2 + \cdots + \hat{x}_u h_{nu} \mu_u = f_n(\hat{x}) \end{aligned} \quad (4)$$

where n corresponds to the number of observations b in the DR image (and corrected observations \hat{b}), \hat{r} are the residuals, i.e. the observation errors, u equals the quantity of attenuation coefficients μ and h gives the occurrence of a certain attenuation coefficient on the respective ray through the CT (from the histograms of n different rays). \hat{x} are unknown factors to the attenuation values.

We solve for factors \hat{x} which modify the attenuations and adapt the CT transfer function to obtain observed intensities b . As negative values of \hat{x} (negative attenuations) do not make physical sense, the solutions are constrained to $\hat{x} \geq 0$ with $i = 1 \dots u$.

Non Negative Linear Least-Squares algorithm of Lawson & Hanson [7] and Singular Value Decomposition are used to solve for for the constrained \hat{x} .

2.3 Registration

Remaining errors \hat{r} are computed by comparing the simulated intensities computed with optimized transfer function to observed DR intensities. Downhill simplex minimization of the root mean square (RMS) of residuals \hat{r} gives the six transformations of the rigid registration of the CT to the X-ray image(s). In each step of the optimization process, the radiometric optimization is repeated and a new set of residuals is generated from (4).

3 Results & Discussion

Our approach was tested using different high resolution CT scans (2 anatomical phantoms of human heads and 1 human pelvis with approximately 1 mm voxel diagonal) and X-ray images acquired in a treatment machine. Image acquisitions with different X-ray tube settings were performed, including kilo-voltage peak settings ranging from 40 to 140 kVp and tube current settings ranging from 80 to 640 mA. In Fig. 2 an X-ray image with a low kVp setting is shown (a). In the center the DRR can be seen that was rendered with the initial X-ray transfer function. With respect to the X-ray image it varies profoundly in its radiometric

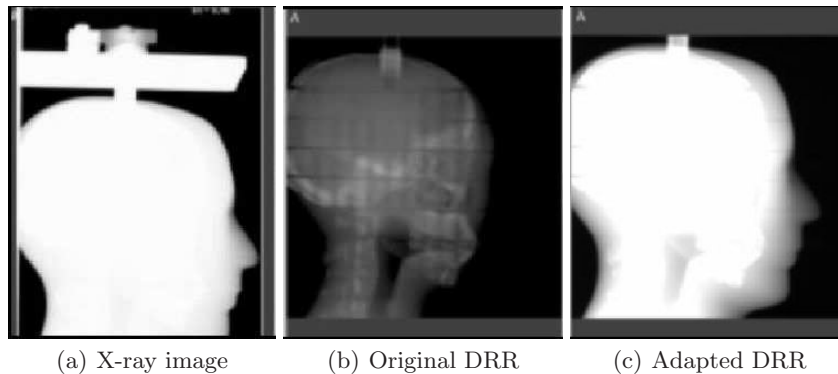


Fig. 2. X-ray image of an anatomical head phantom acquired with the X-ray tube at 40 kVp, 100 mA and 100 ms (a); original DRR image (b); DRR image after several automatic adaptations of the attenuation coefficients during the registration process (c).

properties. In contrast, the DRR on the right, which was generated with the automatically adapted transfer function, looks similar to the DR.

Numerous comparisons of alignment computations using mutual information (MI) versus the proposed TFIR approach were performed. In most cases the target registration accuracy of TFIR was similar to the MI approach (approx. $\frac{1}{2}$ ·DRR Resolution). Tests with DRs of low Signal to Noise Ratio and images containing features (e.g. head fixation device or metallic markers) not visible in the CT show that TFIR is more intolerant regarding those types of image degradation. Additionally, the computation time of TFIR is increased by approximately factor 8 with respect to the MI approach. Nevertheless, results also show that if the X-ray tube voltage is in a very high or low range, automatic transfer function optimization during registration can help to increase registration reliability. Providing higher tolerance against low kVp settings, the TFIR approach could allow reducing the dose delivered to the patient during the X-ray imaging in IGRT.

References

1. Birkfellner W, Wirth J, Burgstaller W, et al. A faster method for 3D/2D medical image registration: A simulation study. *Int J Radiat Oncol Biol Phys.* 2003;48:2665–79.
2. Russakoff DB, Rohlfing T, Mori K, et al. Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D-3D image registration. *IEEE Trans Med Imaging.* 2005;24(11):1441–54.
3. Seibert JA. X-Ray imaging physics for nuclear medicine technologists: Part I: Basic principles of X-Ray production. *J Nucl Med Technol.* 2004;32(3):139–47.
4. Maes F, Collignon A, Vandermeulen D, et al. Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging.* 1997;16(2):187–98.
5. Pluim J, Maintz J, Viergever M. Mutual information based registration of medical images: A survey. *IEEE Trans Med Imaging.* 2003;22(8):986–1004.
6. Selby BP, Sakas G, Stilla U, et al. Reconstruction and registration of multispectral x-ray images for reliable alignment correction in radiation treatment devices. *Proc SPIE.* 2008;6914:2U.1–2U.9.
7. Lawson CL, Hanson RJ. *Solving Least Squares Problems.* Prentice-Hall, Englewood Cliffs, NJ; 1974.

Vorwissensbasierte NFFT zur CT-Metallartefaktreduktion

Bärbel Kratz, May Oehler, Thorsten M. Buzug

Institut für Medizintechnik, Universität zu Lübeck
kratz@imt.uni-luebeck.de

Kurzfassung. Metallobjekte können bei einer CT-Aufnahme zu Artefakten im rekonstruierten Bild führen. Um diesen Einfluss zu reduzieren und die resultierende Bildqualität zu erhöhen, ist eine Artefaktreduktion möglich. In dieser Arbeit wird ein Verfahren basierend auf nicht-äquidistanten Fouriertransformationen (NFFT) als Reduktionsmöglichkeit betrachtet. Durch geeignete Dämpfung dieser Transformation kann Vorwissen einbezogen werden. In Abhängigkeit von der Korrektheit dieser Informationen, kann die resultierende Bildqualität stark variieren. Es werden verschiedene Möglichkeiten der Vorwissensintegration betrachtet und mit NFFT-Interpolationsergebnissen ohne Verwendung dieser a-priori Informationen verglichen.

1 Einleitung

Die Computertomographie (CT) ist eines der am weitesten verbreiteten bildgebenden Verfahren in der medizinischen Diagnostik. Metallobjekte führen bei einer CT-Aufnahme zu starken Inkonsistenzen in den aufgenommenen Rohdaten, wodurch während der Bildrekonstruktion sternförmige Artefakte entstehen. Um die Bildqualität zu erhöhen und damit eine korrekte Diagnose zu unterstützen, kann eine Metallartefaktreduktion (MAR) durchgeführt werden. In den vergangenen drei Jahrzehnten wurden viele MAR-Methoden entwickelt, wobei das Ersetzen der inkonsistenten Rohdaten eine weit verbreitete Strategie ist. Die Datenneubestimmung kann auf verschiedene Arten vorgenommen werden, beispielsweise durch eine Interpolation [1] oder durch nur teilweises Einbeziehen der metallbeeinflussten Rohdaten [2]. Trotz dieser neu ermittelten Daten beinhaltet der Datensatz je nach Qualität der Neubestimmung weiterhin Inkonsistenzen, die zu neuen Artefakten im rekonstruierten Bild führen. Da die überwiegend verwendete gefilterte Rückprojektion (FBP) als Rekonstruktionsschritt sehr sensibel auf diese residualen Inkonsistenzen reagiert, kann daher zusätzlich zur Datenneubestimmung eine alternative Rekonstruktionsmethode in Betracht gezogen werden.

In dieser Arbeit wird ein Interpolationsansatz basierend auf nichtäquidistanten Fouriertransformationen (NFFT) [3] verwendet, wie er in [4] zur Artefaktreduktion präsentiert wurde. Dieses Verfahren ist auf verschiedene Arten auf die konkrete Fragestellung optimierbar. Je nach Vorliegen der zu bearbeitenden Daten kann zum Beispiel eine angepasste Interpolationsdimension gewählt werden.

Dadurch werden nicht nur parallele Projektionspfade mit in den Interpolationsschritt einbezogen.

Des Weiteren ist die Verwendung von Vorwissen möglich, sofern dieses gegeben ist, was zu einer deutlich verbesserten Bildqualität führen kann. Zwei verschiedene Integrationsmöglichkeiten von a-priori Informationen werden folgend vorgestellt und abschließend mit NFFT-Interpolationsergebnissen ohne Vorwissen verglichen.

2 Material und Methoden

Die MAR-Methoden werden auf zwei verschiedene CT-Datensätze angewendet. Zur Evaluierung wurden zwei Aufnahmen eines anthropomorphen Torsophantoms (CIRS Inc.) durchgeführt, mit und ohne Metallobjekte. So ergibt sich zu einem metallbeeinflussten Datensatz (Abb. 1(a) im Radonraum und 1(b) im Bildbereich) ein Referenzdatensatz, mit dem zur Bewertung der MAR-Ergebnisse ein Vergleich durchgeführt werden kann. Anhand eines Hüftimplantat-Datensatzes (Abb. 1(c) im Radonraum und 1(d) im Bildbereich) soll außerdem ein Beispiel der MAR-Anwendung für klinische Daten gegeben werden.

Seien $p(\mathbf{r})$, $\mathbf{r} = (\gamma, \xi)$ die erfassten Radondaten an den Positionen \mathbf{r} mit Projektionswinkel γ und Detektorposition ξ . Um eine Neubestimmung der metallbeeinflussten Daten vorzunehmen, wird zuvor eine Separierung von metallbeeinflussten und nicht beeinflussten Daten benötigt. Durch eine vorläufige FBP-Rekonstruktion und eine Schwellwertanwendung kann eine Maske gewonnen werden, die transformiert in den Radonraum eine Trennung der Metalldaten \bar{R} an den Stellen $\bar{\mathbf{r}}_n$ und den restlichen Daten R' an den Positionen \mathbf{r}'_j ermöglicht. Nach diesem Maskierungsschritt können die entfernten Werte auf Basis der verbliebenen Daten neu bestimmt werden.

Wie in [4] vorgestellt, kann nach der Maskierung im Radonraum - basierend auf den verbleibenden, nichtäquidistant verteilten Stützpunkten \bar{R} - eine Dateneubestimmung durch eine zweidimensionale NFFT [3] durchgeführt werden, welche in Matrix-Vektorschreibweise durch

$$\mathbf{p} = \mathbf{B}\hat{\mathbf{p}}, \text{ mit } \mathbf{p} = (p(\mathbf{r}'_j))_{j=0}^{R'-1}, \mathbf{B} = (e^{2\pi i \mathbf{k}_\kappa^T \mathbf{r}'_j})_{\kappa,j=0}^{R-1, R'-1}, \hat{\mathbf{p}} = (\hat{p}_\kappa)_{\kappa=0}^{R-1} \quad (1)$$

gegeben ist. Die Variablen $\mathbf{k}_\kappa \in [-\frac{R_1}{2}, \frac{R_1}{2}] \times \dots \times [-\frac{R_d}{2}, \frac{R_d}{2}]$ entsprechen äquidistanten Frequenzen, $R = R_1 \dots R_d$ definiert die Datensatzgröße mit den Dimen-

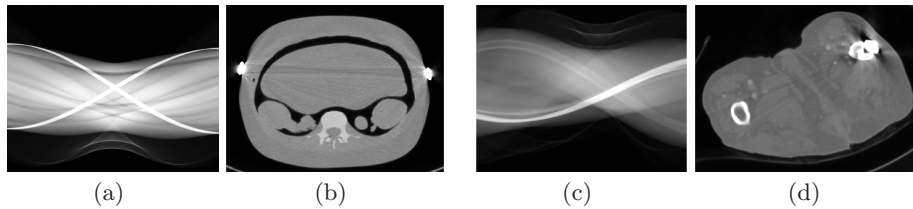


Abb. 1. Phantom und Hüftdatensatz im Radonraum (a,c) und im Bildbereich (b,d).

sionsgrößen R_1, \dots, R_d , R' entspricht der gegebenen Anzahl von Stützpunkten an den Positionen \mathbf{r}'_j und \hat{p}_κ beschreibt die zugehörigen Fourierkoeffizienten. Durch Auswertung von (1) an den Positionen $\bar{\mathbf{r}}_n$ ergibt sich eine Interpolation im Maskenbereich. Da im Rahmen dieser Anwendung $R' < R$ gilt, ist (1) unterbestimmt und kann durch das gedämpfte Minimierungsproblem

$$\|\hat{\mathbf{p}}\|_{\hat{\mathbf{w}}}^2 = \sum_{\kappa=0}^{R-1} \frac{|\hat{p}_\kappa|^2}{\hat{w}_\kappa} \xrightarrow{\hat{\mathbf{p}}} \min \quad \text{mit} \quad \mathbf{p} = \mathbf{B}\hat{\mathbf{p}}, \quad (2)$$

gelöst werden [5]. Die Dämpfungsfaktoren $\hat{\mathbf{w}} = (\hat{w}_\kappa)_{\kappa=0}^{R-1}$ können, falls vorhanden, Vorwissen in die Koeffizientenberechnung einfließen lassen. Liegt kein Vorwissen vor, so ist eine $\hat{\mathbf{w}}$ -Belegung anhand von vorangegangenen Interpolationsergebnissen möglich, um eine Richtung der gesuchten Lösung approximativ vorzugeben. Im Rahmen dieser Arbeit wurden dazu beispielhaft eine lineare und eine kubische Interpolation betrachtet.

Ein Ansatz zur Wertneuzuweisung im Maskenbereich wurde von [2] inspiriert. Hierbei wird nach einer Rekonstruktion als Zwischenschritt eine neue Wertzuweisung der Metallbereiche im Bildraum vorgenommen. Dies kann auf verschiedene Arten durchgeführt werden. Eine Möglichkeit liefert ein Einbeziehen der originalen Werte im Maskenbereich des rekonstruierten Bildes, jedoch mit reduzierter Intensität. So sollen die im Verhältnis zum restlichen Datensatz sehr hohen Metallwerte keinen negativen Einfluss auf die spätere Rekonstruktion haben. Hier wurde eine prozentuale Gewichtung von 50 % gewählt, um ein erneutes Vorkommen der Metallartefakte zu vermeiden.

Alternativ dazu kann die Metallnachbarschaft im rekonstruierten Bild betrachtet werden. Es ist wahrscheinlich, dass im Metallbereich ohne Metallobjekte ähnliche Werte vorliegen würden, wie in der näheren Umgebung. Daher bildet eine Möglichkeit die Verwendung des Medianwertes der Nachbarschaft des Maskenbereichs. Im Rahmen dieser Arbeit wurden die Ergebnisse für eine 8-Nachbarschaft verwendet. Bei einer zu groß gewählten Nachbarschaftsbetrachtung erhöht sich die Gefahr, in andere Strukturbereiche zu gelangen und somit einen abweichenden Wert für den Metallbereich zu ermitteln.

Eine Transformation zurück in den Radonraum liefert für beide Ansätze eine Belegung im Maskenbereich für die Dämpfungsfaktoren in Gl. (2).

Da die FBP sehr sensitiv auf Inkonsistenzen innerhalb der Rohdaten reagiert (ein Vergleich wurde in [6] durchgeführt), wurde hier zur Rekonstruktion der Interpolationsergebnisse ein gewichteter Maximum-Likelihood-Expectation-Maximization-Ansatz (λ -MLEM) [7] verwendet. Bei der Gewichtung handelt es sich um einen Vertrauensfaktor λ , der je nach Art der Werte zwischen 0 und 1 variiert werden kann. Insbesondere bei interpolierten Daten ergibt sich somit die Möglichkeit, durch eine reduzierte Gewichtung diese neuen Werte nur teilweise in die Rekonstruktion einfließen zu lassen. Dies liefert einen optimalen Kompromiss zwischen interpolierten Werten mit residualen Fehlern und ein Nichteinbeziehen der gesamten Daten im Maskenbereich.

3 Ergebnisse

Abb. 2 zeigt die jeweils besten rekonstruierten MAR-Ergebnisse mit variierender Dämpfung, ohne und mit der Integration von a-priori Informationen. Die kubisch gedämpfte NFFT (Abb. 2(a)) lieferte das beste Ergebnis ohne die Verwendung von Vorwissen. Bei der Einbeziehung des Median der Nachbarschaft ergab sich die größte Qualitätsverbesserung aller verwendeten Ansätze (Abb. 2(b)).

Alle Ansätze reduzieren die Anzahl von Metallartefakten, wobei jedoch neue Bildfehler entstehen. Die Ursache dafür liegt im Interpolationsschritt, der zu einem gewissen Grade erneut fehlerhafte Werte einbringt. Die Anzahl und Intensität dieser neuen Strukturen kann durch die NFFT in Kombination mit Vorwissen jedoch deutlich reduziert werden. Ein numerischer Vergleich basierend auf den relativen Abweichungen zum Referenzdatensatz aller betrachteten Verfahren ist in Abb. 3 dargestellt. Die Kurven entsprechen dabei einem variierenden Vertrauensfaktor λ für den Rekonstruktionsschritt. Für jedes Verfahren ist markiert, welche λ -Belegung zu einer minimalen Referenzabweichung führt. Alle Interpolationen weisen eine gute Annäherung an die Referenz auf, wobei die Median-gedämpfte NFFT die beste relative Fehlerreduktion liefert.

Angewendet auf den klinischen Hüftdatensatz ergeben sich die Bilder, welche in Abb. 2(c, d) dargestellt sind. Hier wird deutlich, dass unter Hinzunahme von

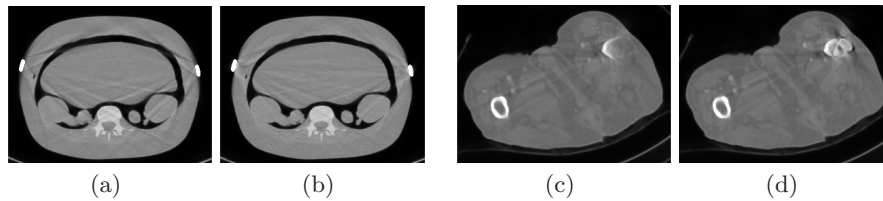


Abb. 2. Kubisch- und Median-gedämpfte NFFT Interpolation des Phantoms (a,b) und des Hüftdatensatzes (c,d).

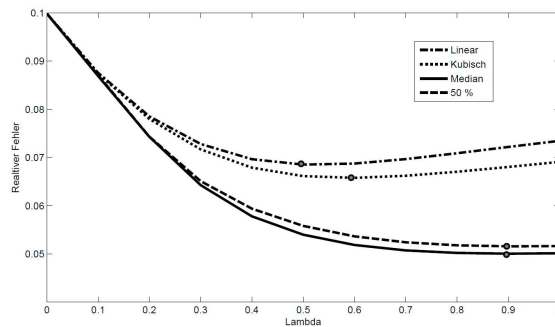


Abb. 3. Relativer Fehler $\|(\mathbf{Bild}_{\text{Referenz}} - \mathbf{Bild}_{\text{MAR}})\|_2 / \|\mathbf{Bild}_{\text{Referenz}}\|_2$ (ohne Betrachtung des Metallbereiches) der einzelnen Ansätze in Abhängigkeit von verschiedenen Werten λ während der Rekonstruktion.

Vorwissen während des Interpolationsschrittes deutlich mehr Strukturinformationen im Metallbereich erhalten und neu gewonnen werden.

4 Diskussion und Ausblick

Der NFFT-Ansatz besitzt einige Vorteile gegenüber herkömmlichen Interpolationsverfahren. Neben der Möglichkeit, die Interpolationsdimension an den gegebenen Datensatz anzupassen, kann Vorwissen in die Transformation einbezogen werden. In Kombination mit einer λ -MLEM-Rekonstruktion zeigt sich, dass mit dem Einbringen von korrektem Vorwissen in den Interpolationsschritt der Vertrauensfaktor während der Rekonstruktion deutlich erhöht werden kann. Dies bestätigt, dass die Interpolationsergebnisse weniger residuale Fehler beinhalten und somit zu einer besseren Bildqualität führen.

Zukünftig soll die Integration von Vorwissen in die NFFT-Interpolation weiter optimiert werden. Im Rahmen dieser Arbeit wurde das prozentuale Einbeziehen im Objektraum auf die Hälfte des Originalwertes gesetzt, um das erneute Aufkommen von Metallartefakten zu vermeiden. Eine Schwelle, bis zu der diese Werte maximal positiv in die spätere Rekonstruktion des Bildes eingehen, ist noch genauer zu evaluieren. Des Weiteren sollten ebenfalls unterschiedliche Nachbarschaftsdefinitionen auf ihre Vor- und Nachteile hin überprüft werden.

Literaturverzeichnis

1. Kalender WA, Hebel R, Ebersberger J. Reduction of CT artifacts caused by metallic implants. *Radiology*. 1987;164:576–7.
2. Lemmens C, Faul D, Nuyts J. Suppression of Metal Artifacts in CT Using a Reconstruction Procedure that combines map and projection completion. *IEEE Trans Med Imaging*. 2009;28(2):250–60.
3. Potts D, Steidl G. New fourier reconstruction algorithms for computerized tomography. *Proc SPIE*. 2000;4119:13–23.
4. Kratz B, Knopp T, Müller J, et al. Comparison of nonequispaced fourier transform and polynomial based metal artifact reduction methods in computed tomography. In: *Proc BVM*; 2008. p. 21–5.
5. Kunis S, Potts D. Stability results for scattered data interpolation by trigonometric polynomials. *SIAM J Sci Comput*. 2007;29:1403–19.
6. Kratz B, Oehler M, Buzug TM. On limitations of 1d interpolation-based metal artefact reduction approaches. A comparison of fbp versus mlem. In: *Proc World Congr Med Phys Biomed Eng*. vol. 25/II; 2009. p. 398–401.
7. Oehler M, Buzug TM. Statistical image reconstruction for inconsistent CT projection data. *Methods Inf Med*. 2007;46:261–9.

Parahistogramme innerhalb eines dreidimensionalen Interaktionsraumes

Albert Pritzkau, Dirk Bartz

Universität Leipzig, ICCAS/VCM
albert.pritzkau@medizin.uni-leipzig.de

Kurzfassung. Im Kontext der Genexpressionsanalyse haben sich parallele Koordinaten (PK) als ein hilfreiches Werkzeug erwiesen. Ein hochdimensionaler Datensatz kann auf diese Weise in einem zweidimensionalen Datenraum dargestellt werden. Dabei wird die Anzahl der darstellbaren Dimensionen lediglich durch den horizontal verfügbaren Platz begrenzt und ermöglicht ein relativ leichtes Erkennen inhärenter Zusammenhänge innerhalb eines Kontextes. Neben den offensichtlichen Vorteilen sind jedoch auch die damit verbundenen Beschränkungen zu berücksichtigen. Durch Überlagerung einzelner Linien kann diese Darstellung relativ schnell überzeichnet werden und verliert in diesen Bereichen an Informationsgehalt. Darüber hinaus stellt sich die Detektion von Korrelationen über mehrere Achsen als eine relativ schwierige Aufgabe dar. Um die Frequenz von Kantenzügen entlang der Koordinatenachsen zu erhalten, werden zur graphischen Darstellung Histogramme herangezogen, welche jeweils an den Koordinatenachsen ansetzen. Ohne den Kontext der PK-Darstellung zu verlassen, werden quantitative Aussagen über die Verdeckung in einem bestimmten Wertintervall ermöglicht.

1 Einleitung

Mit Einführung verschiedener Genom- und Proteom-Technologien wird die gleichzeitige Messung von vielen Hunderten oder Tausenden von biologischen Messgrößen ermöglicht. Diese ergeben sich aus unterschiedlichen medizinischen und biologischen Problemstellungen. Die Etablierung geeigneter Algorithmen zur Analyse dieser meist hochdimensionalen Daten stellt ein entscheidendes Kriterium dar. Die Herausforderung der Genexpressionsanalyse besteht beispielsweise darin, verlässliche Aussagen über die Aktivität von Genen unter unterschiedlichen Umständen zu treffen. Microarrays oder DNA-Chips dienen der Bestimmung relativer Änderungen der Genexpression. Das Ergebnis solcher Experimente setzt sich jeweils aus einer Liste der ermittelten Gensequenzen mit den dazugehörigen Expressionswerten dar. Mehrere Experimente ergänzen sich zu einer Datenmatrix dessen Spalten jeweils einem Experiment zuzuordnen sind. Zur Gewinnung aussagekräftiger statistischer und biologischer Daten greift man zunehmend auch auf grafische sowie andere Methoden zur Visualisierung solcher komplexer Datensätze zurück [1]. Bei der visuellen Analyse müssen dabei die hochdimensionale Daten auf niedrigdimensionale Sichten projiziert werden.

Bezüglich der visuellen Datenexploration von hochdimensionalen Daten hat sich die Darstellung der parallelen Koordinaten [2] (PK) als ein sehr hilfreiches Werkzeug erwiesen. Diese Form der Visualisierung veranschaulicht einen k -dimensionalen Datensatz anhand von zwei Darstellungsdimensionen. Sie besteht aus k parallelen und typischerweise äquidistanten Achsen. Diese repräsentieren jeweils den Wertebereich einer Dimension. In dem oben beschriebenen Ausgangsdatsatz ist eine Dimension jeweils einem Experiment zuzuordnen. Der dazugehörigen Wertebereich umfasst die minimalen und maximalen Expressionswerte. Generell werden benachbarte Achsen gleich orientiert. Über eine Nulllinie können die einzelnen Achsen zusätzlich miteinander verbunden werden. Zur Repräsentation werden die Datenwerte jeweils einer Zeile auf den entsprechenden Achsen abgetragen und ergeben einen Kantenzug. Das Ergebnis kann als Grundlage anschließender Analysen genutzt werden.

2 Material und Methoden

Die Darstellung der PK eignet sich ausgezeichnet zur Identifikation von Korrelationen zwischen Attributen benachbarter Achsen. Darüber hinaus sind auch Werteverteilungen oder Ausreißer gut erkennbar. Jedoch stellt sich in Bereichen hoher Linienkonzentrationen, die sich vor allem durch ähnliche oder mehrfach auftretende Attributwerte ergeben, die Interpretation der Darstellung als relativ schwierig dar. Dieses Problem wurde bereits durch eine Reihe unterschiedlicher Ansätze adressiert. Zur Darstellung einer Häufigkeitsverteilung entlang des Wertebereichs einer Koordinatenachse wurden beispielsweise Histogramme zur Integration in die PK-Darstellung vorgeschlagen [3]. Diese Kombination folgt generell dem weithin bekannten Konzept der Linked-Views, und findet auch in aktuelleren Beiträgen wie [4, 5] besondere Beachtung.

In der vorliegenden Arbeit wird die Einbettung der traditionell zweidimensionalen PK in einen dreidimensionalen Interaktionsraum vorgeschlagen. Es existieren bereits eine Reihe unterschiedlicher Ansätze, die hinzu gewonnene Dimension vorteilhaft zu nutzen. Beispielsweise erweitern [6, 7] die generelle Struktur der PK, um die Darstellung auf diese Weise zu entzerren. Im Gegensatz dazu wird in unserem Fall die traditionelle PK-Darstellung innerhalb des gegebenen Interaktionsraumes koplanar zur xy -Ebene angeordnet. Die räumliche Dimension wird dazu genutzt, eine Histogrammdarstellung zu integrieren. Sie ermöglicht anhand der angereicherten Häufigkeitsverteilung der Kantenzüge entlang der Koordinatenachsen eine differenzierte Interpretation kritischer Bereiche. Wie in Abb. 1 dargestellt, verlaufen die Histogrammsäulen, welche an den Koordinatenachsen ansetzen, orthogonal zur PK-Darstellung in z -Richtung des Interaktionsraumes. Die Höhe einer Säule repräsentiert dabei jeweils Frequenz von Kantenzügen in dem zugehörigen Intervall. Ist nun der Interaktionsraum entlang der Blickrichtung des Beobachters ausgerichtet, werden die Histogrammsäulen zunächst auf die jeweiligen Koordinatenachsen projiziert und sind für den Benutzer nicht sichtbar. Durch einfache Navigationsinteraktionen kann jedoch der Interaktionsraum in eine gewünschte Position rotiert werden und eröffnet damit den Blick

auf die Histogrammdarstellung ohne den visuellen Kontakt zu den PK zu unterbrechen.

3 Ergebnisse

Die Einbettung der PK-Darstellung in einen drei dimensional Interaktionsraum ermöglicht die Integration von weiteren Datenattributen. Durch Rotationen des Interaktionsraumes wird die Sichtbarkeit von diesen Zusatzinformationen - hier die Histogrammdarstellung - beeinflusst. Die Ausrichtung der Darstellung in räumliche Dimensionen ermöglicht eine visuelle Gewichtung einzelner Bestandteile des Interaktionsraumes, die je nach Bedarf von der vollständigen Sichtbarkeit bis hin zur Verdeckung reicht.

Der visuelle Vergleich einzelner Histogramme kann zudem als Ansatzpunkt weiterer Analysen dienen. So lassen sich beispielsweise Korrelationen zwischen Achsen mit ähnlicher Häufigkeitsverteilung vermuten. Zur differenzierten Untersuchung können zusätzlich die einzelnen Histogrammbalken markiert werden. Alle Linienzüge, die durch das entsprechende Intervall verlaufen, werden farblich hervorgehoben (Abb. 2). Anhand dieser Markierung ist es möglich Korrelationen über mehrere Dimensionen hinweg zu erörtern. Zur automatischen Selektion der einzelner Histogrammbalken wird eine Ebene eingesetzt werden, die parallel zur Ebene der parallelen Koordinaten angeordnet ist. Über einen Schieberegler kann die Höhe der Selektionsebene gesteuert, welche eine bestimmte Auftrittshäufigkeit repräsentiert. Alle Histogrammsäulen, die diese Ebene schneiden und damit den gegebenen Häufigkeitswert überschreiten, werden markiert. Diese Selektion berücksichtigt generell die Histogramme der gesamten PK-Darstellung. Jedoch kann die Auswahl auch auf eine Dimension beschränkt werden.

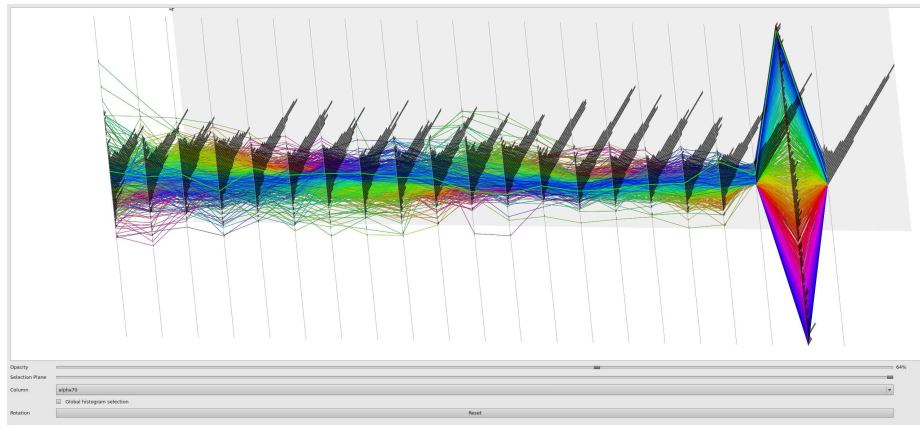


Abb. 1. Angereicherte Darstellung der Zeitreihe eines Zellzyklus-Experiment an Bäckerhefe (*Saccharomyces cerevisiae* [8, 1]) erweitert durch Histogramme.

4 Diskussion

Im Gegensatz zu den bekannten Verfahren, wird die Visualisierung der PK grundsätzlich nicht durch die Darstellung zusätzlicher Attribute überlagert. Jedoch ist der Benutzer in der Lage durch entsprechende Interaktionen den Fokus individuell so zu manipulieren, dass die erforderlichen Informationen für sichtbar werden. Anstelle der Histogrammdarstellung sind auch weitere Darstellungsformen denkbar, die die räumliche Dimension ausnutzen. Beispielsweise verfolgen Parallel Sets [9] einen ähnlichen Ansatz zur Analyse. Auch hier wird der Wertebereich einer Achse in eine Folge von Intervallen unterteilt und die Häufigkeit der betroffenen Linienzüge aufsummiert. Im Gegensatz zu Histogrammen rückt hier jedoch die Verteilung der Linienzüge zwischen den benachbarten Achsen in den Fokus der Betrachtung. Diese werden durch Bänder repräsentiert, die jeweils zwei Intervalle benachbarter Koordinatenachsen miteinander verbindet. In Zukunft soll die Integration dieser Methode in den vorhandenen Interaktionsraum evaluiert werden, um so vielleicht noch besser Korrelationen über mehrere Achsen hinweg verfolgen zu können.

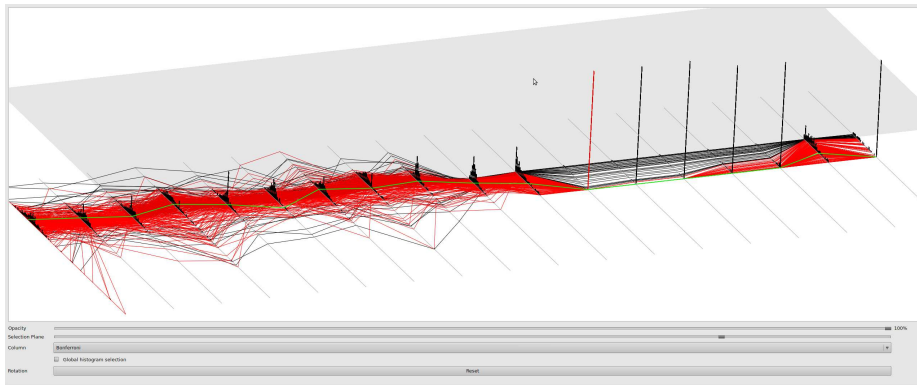


Abb. 2. Angereicherte Darstellung einer Expressionsstudie zur Auswirkung von sportlicher Anstrengung auf das Immunsystem [10, 1] erweitert durch eine Selektionsebene. Histogrammsäulen einer Koordinatenschase, welche die Selektionsebene schneiden, werden selektiert. Die Selektionsebene ist hier halbtransparent dargestellt.

Literaturverzeichnis

1. Dietzsch J, Heinrich J, Nieselt K, et al. SpRay: A visual analytics approach for gene expression data. In: Proc IEEE Symp Vis Anal Sci Technol; 2009.
2. Inselberg A, Dimsdale B. Parallel coordinates: a tool for visualizing multi-dimensional geometry. In: Proc IEEE Vis; 1990. p. 361–78.
3. Ong HL, Lee HY. WINVIZ: a visual data analysis tool. Comput Graph. 1996;20(1):83–4.

4. Hauser H, Ledermann F, Doleisch H. Angular brushing of extended parallel coordinates. In: Proc IEEE Symp Inform Vis; 2002. p. 127–30.
5. McDonnell KT, Mueller K. Illustrative parallel coordinates. Comput Graph Forum. 2008;27(3):1031–8.
6. Johansson J, Cooper M, Jern M. 3-Dimensional display for clustered multi-relational parallel coordinates. In: Proc IEEE Int Conf Inform Vis; 2005. p. 188–93.
7. Fanea E, Carpendale MST, Isenberg T. An interactive 3D integration of parallel coordinates and star glyphs. In: Proc IEEE Symp Inform Vis; 2005. p. 149–56.
8. Spellman PT, Sherlock G, Zhang MQ, et al. Comprehensive identification of cell cycle-regulated genes of the yeast *saccharomyces cerevisiae* by microarray hybridization. *Molecular Biology of the Cell*. 1998;9(12):3273–97.
9. Kosara R, Bendix F, Hauser H. Parallel sets: Interactive exploration and visual analysis of categorical data. *IEEE Trans Vis Comput Graph*. 2006;12(4):558–68.
10. Zieker D, Fehrenbach E, Dietzsch J, et al. cDNA microarray analysis reveals novel candidate genes expressed in human peripheral blood following exhaustive exercise. *Physiol Genomics*. 2005;23(3):287–94.

Acceleration of the Fully Automatic Branch Labeling of Voxel Vessel Structures

Lei Chen¹, Jan Bruijns², Bart M. Haar ter Romeny¹

¹Biomedische Technologie, Technische Universiteit Eindhoven

²Philips Research Eindhoven, The Netherlands

chen@isip.uni-luebeck.de

Abstract. For diagnosing a stenosis or an aneurysm, the shape parameters of the diseased vessel parts are needed by physicians. Therefore, a fully-automatic extraction of this shape from a volume representation has been developed. This paper analyses the blood vessel branch labeling acceleration algorithms and proposes two improved methods. The first one is called the surface wave propagation method which restricts the wave moving along the blood vessel surface. The second one replaces the thinning algorithm with the surface propagation to extract the center lines and the bifurcations of the blood vessels. Experiment results show that two improved methods can decrease the computation time and keep the labeling accuracy.

1 Problem

3D Rotational Angiography system (3DRA) enables physicians to capture detailed 3D images of a patient vascular structure, which leads to faster and more accurate diagnosis and treatment. It is increasingly used to diagnose brain aneurysms, Gauvrit et al. [1], and to decide on the optimal treatment modality. Volume representations of blood vessels acquired by 3DRA after injection with a contrast agent have a clear distinction in gray values between tissue and vessel.

For optimal treatment, physicians need to know the shape parameters of the diseased vessels parts, Bruijns [2]. A method for fully-automatic labeling of blood vessels voxels has been developed by Bruijns[3]. This method starts with the detection of the extremities of the vessel by a wave propagation method by Zahlten et al.[4] in the segmented volume. The second step, a thinning approach is used to peel the blood vessels in a number of iterations. At last, the vessel vertices are labeled with a unique number for each vessel branch.

In this paper, we present two acceleration algorithms. First, we implemented a wave propagation confined to move along the blood vessel surface only (SWP). The second acceleration method replaces the thinning algorithm with the SWP to extract the center lines and bifurcations of the blood vessels.

2 Related Work On Branch Labeling

The graph structure can be generated by various skeletonization algorithms, which represents the topology of the blood vessels. Skeletonization methods

based on morphological thinning are presented [5]. The topology conditions resulted occasionally in small cycles, which cannot be presented in the segmented voxel volume in their thinning method. A single skeleton is directly generated from the gray value volume [6]. But this method cannot be used for separated components. Since the vessel voxels cannot be labeled according to the branch or center region they belong to. A system for reconstruction of a vessel tree is created by a wave propagation (WP) [4]. The WP method labels the original gray value volume using an appropriate threshold and generates the corresponding vessel graph. However this method is not accurate enough at the bifurcations (Fig. 1(a, b)).

A method for robust mapping of bifurcating vessels is presented [7]. But a 3D triangle surface model of the vessel boundaries is needed as input for their computations. A thinning method is used after a wave propagation [8]. A segmented voxel volume with the extremities is set as the input data, the resulting skeleton of branches and bifurcations is a better approximation of the vessel graph than the method of Zahlten et al.. However the thinning is a time consuming approach for the bifurcation position correction.

3 Acceleration Methods of Branch Labeling

The branch labeling method consists of five main parts, and the first two of them are time consuming [8]. In this paper, the volume wave propagation (VWP) is replaced by the surface wave propagation (SWP). Then the first two steps are combined to one, the thinning processing is excluded. During the extremity detection, the center lines and bifurcations are extracted by the SWP.

3.1 Surface Wave Propagation

For the VWP, a wave consists of voxels which mostly form a cross section of the blood vessel. When the wave stops moving, a voxel could be found from the final wave, its distance to the seed point is the farthest. This voxel can be saved into an extremity set as an extremity. The extremities detection of the vessels inside

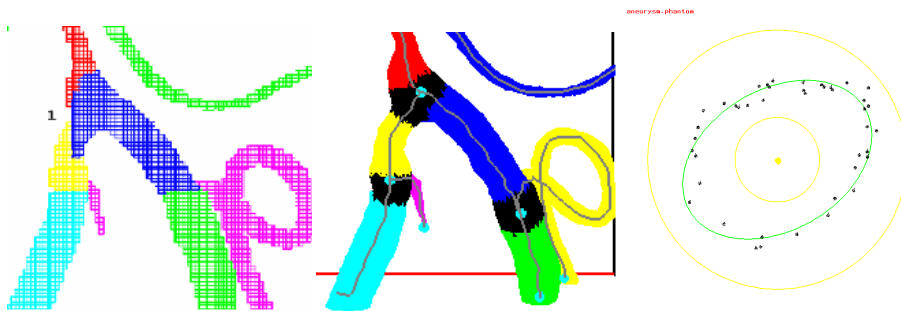


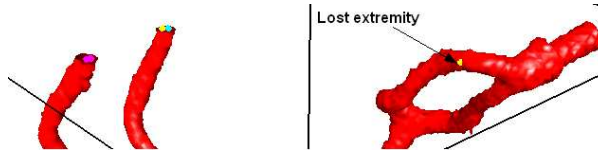
Fig. 1. Left: wrong bifurcation; middle: bifurcation position; right: fitted ellipse.

the volume is implemented by means of a wave propagation. The computing time depends mainly on the generation of a new wave from an old wave. When the inner vessel voxels can be ignored by giving them a special label, the waves will be much smaller because only the surface vessel voxels are included. In this case the waves will propagate along the vessel surface.

The starting point is a segmented volume with tissue voxels, aneurysm voxels and normal vessel voxels [9]. The initial wave was filled by the surface voxels after the selection of a seed point. During the double WG [8], all the neighbors of the old wave are found, in which only the surface voxels would be put into the new wave.

In the SWP, the center point of the stop wave could be set as the extremity for this branch. A false extremity would be detected by the original method when a cycle presents in the segmented voxel volume. In such a case the detected extremity is not at the branch tip. A yellow point labeled with "lost extremity" can be seen on the cycle in Fig. 2. A simple solution is to check whether all neighbors of the stop wave are tissue voxels. If any neighbor voxel has different branch number from the stop wave, it will be considered as a false extremity. In Fig. 2, the yellow and the cyan points indicate the extremities which are detected by the VWP and the SWP respectively; and the pink symbol means the detected results at the same voxel.

Fig. 2. The extremity detection comparison.



3.2 Center Lines Extraction and Bifurcation Detection

During movement of the wave, the center and the normal of a wave are used for the direction adjustment. The plane (a wave) should be orthogonal to the vessel. An oblique plane would give the wrong parameters such as a diameter. This section describes a method of fitting a plane to uncertain 3D points $\{P_i = (x_i, y_i, z_i)^T | i = 1 \dots N\}$. All data points P_i that lie on the plane defined by the normal $n = (A, B, C)$ and the perpendicular distance to the original d . The value ϵ is introduced on the right of equation $Ax_i + By_i + Cz_i - d = \epsilon_i$ standing for the fitting error. The plane normal n gives the homogeneous least square problem, The solution is computed by means of singular value decomposition. In the same way, a least-square ellipse is directly fitted through a number of 3D points in a wave for the visualization. As shown in Fig. 1(c, a) yellow center and a green fitted ellipse indicate a wave.

Generally, bifurcations can be found when the wave is separated to two or more new waves. In this paper, the bifurcations can be positioned at where

Table 1. Statistical result of the volume and the surface wave propagation. For the decreased time, we obtained: min = 24%, max = 82%, mean = 50%, and std = 0.29.

Images	Extremity	Elapsed	False Ex	Lost Ex	Decreased time
ane p	7, 6	0.66, 0.37	1	0	44%
sten-1	22, 22	0.33, 0.18	0	0	45%
sten-p	51, 49	0.29, 0.19	3	1	36%
lane02	55, 53	48.5, 15.3	12	18	68%
lane50	163, 180	7.23, 4.76	46	11	34%
lane52	81, 68	36.4, 11.3	18	5	69%

the ellipses disconnect each other and defined as two main types. When the wave moves from a wider vessel to two or more narrower vessels, the wave could be disconnected. In this case the current wave center is defined as the type I bifurcation position. Conversely, when the wave moves from a narrower vessel to a wider vessel, the location where are the disconnected waves should be defined as the type II bifurcation. In Fig. 3(a), the gray parts indicate the bifurcation type II and these black symbols circle the bifurcation positions.

4 Experiments Results and Validation

The SWP has been applied for fully-automatic branch labeling algorithm [8] to 60 clinical volume datasets. Compare to the VWP, the computation time is decreased at least 24 % and at most 82 %. Parts of the validation result is given in Tab. 1. The size of the first three images is $128 \times 128 \times 128$ and the others are $256 \times 256 \times 256$. The extremity amount, the elapsed time and the amount of

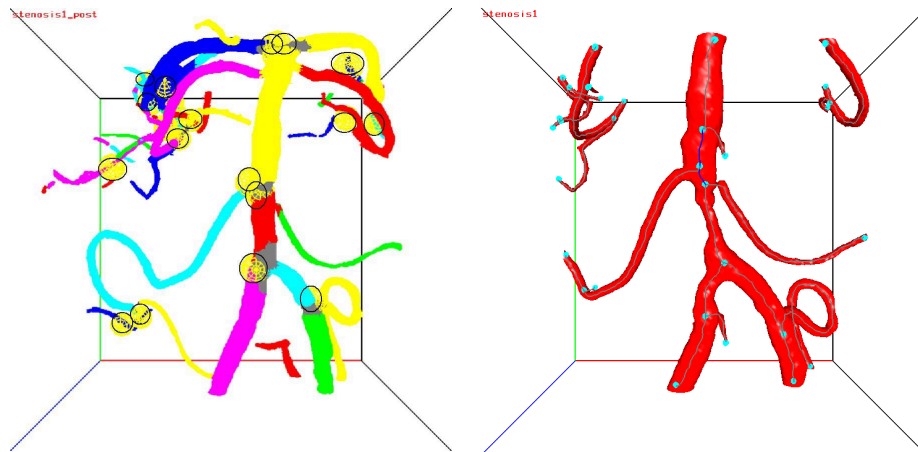


Fig. 3. (a) the bifurcation detection;(b) the labeling result.

the false or lost extremity are given in tab. 1. The two numbers in each grid of the first two columns are calculated with the methods based on the VWP and SWP respectively. There are eight datasets have been manually tested for the bifurcation detection. Almost all bifurcations can be detected, which are indicated as the black symbol in Fig. 3(left). At last the labeling result is shown in Fig. 3(right).

5 Discussion and Conclusions

The following conclusions can be drawn from the results, the figures and the experiences gathered during testing. SWP always gives the correct and visually acceptable extremity detection results. The comparison result with the VWP shows that the detection time is decreased at least 24% and at most 82%. The centerlines and bifurcations extraction by SWP give the correct and visually acceptable results. Based on the extremity set, the center voxels set and the bifurcation set, the vessel graph is created but it needs to be evaluated. Whether the branch labeling result, generated by SWP based on a vessel graph, are suitable for computer assisted diagnosis has not been investigated yet.

References

1. Gauvrit JY. 3D rotational angiography: Use of propeller rotation for the evaluation of intracranial aneurysms. *Am J Neuroradiol.* 2005;26(1):163–5.
2. Bruijns J, Peters FJ, Berretty RPM, et al. A method to detect and mark false branches of a vessel graph. In: *Proc SOIA*; 2006. p. 159–70.
3. Bruijns J. Fully-automatic branch labelling of voxel vessel structures. In: *Proc VMV*; 2001. p. 341–50.
4. Zahlten C, Juergens H, Peitgen HO. Reconstruction of branching blood vessels from CT-data. In: *Proc Vis Sci Comp*; 1994. p. 41–53.
5. Bertrand G, Malandain G. A new characterization of three-dimensional simple points. *Pattern Recognit Lett.* 1994;2(15):169–75.
6. Dokladal P. Grey-Scale Image Segmentation: A Topological Approach [PhD Thesis]. University Marne La Vallee. France; 2000.
7. Antiga L, Steinman DA. Robust and objective decomposition and mapping of bifurcating vessels. *IEEE Trans Med Imaging.* 2004;23(6):704–13.
8. Bruijns J. Fully-Automatic Branch Labelling of Voxels in Vessel Structures. Eindhoven, The Netherlands: Philips Research; 2001. NL-TN 2001/058.
9. Bruijns J, Peters FJ, Berretty RPM, et al. Fully-automatic correction of the erroneous border areas of an aneurysm. In: *Proc BVM*; 2007. p. 293–7.

Photogrammetrische 3D-Vermessung von Organen

Michael Witte¹, Cora Wex¹, Nils Riefenstahl², Bernd Michaelis³,
Stephan Jacob¹, Hans Lippert¹

¹Universitätsklinik für Allgemein-, Viszeral- und Gefäßchirurgie,
Otto-von-Guericke-Universität Magdeburg

²INB Vision AG, Magdeburg

³Institut für Elektronik, Signalverarbeitung und Kommunikationstechnik,
Otto-von-Guericke-Universität Magdeburg

michael.witte.ovgu@gmx.de

Kurzfassung. In der vorliegenden Arbeit wird ein Verfahren vorgestellt, welches die Bestimmung der Oberflächengeometrie von Organen bei einem laparoskopischen Eingriff ermöglicht. Die Aufnahmen erfolgen über ein Stereokamerasystem, bei dem die Objektive der Kameras durch Messendoskope ersetzt wurden. In einem ersten Schritt werden aus einem aufgenommenen Bildpaar unter Verwendung eines Flächenkorrelationsverfahrens dreidimensionale Näherungswerte für die Organoberfläche bestimmt. Mit Hilfe der Flächenkorrelation im Objektbereich nach Michaelis und Albrecht [1] werden diese Näherungswerte anschließend präzisiert. Zur Lösung des Korrespondenzproblems in homogenen Bereichen der Organtextur wird ein stochastisches Rauschmuster auf die Oberfläche projiziert.

1 Einleitung

Im Rahmen des vom Bundesministerium für Wirtschaft und Technologie gefördert VIRTOP-Netzwerkes sollte eine Möglichkeit geschaffen werden, die Ausbildungsbedingungen für operativ tätige Ärzte zu verbessern, bevor diese am realen Patienten operieren. Zu diesem Zweck sollte ein Operationssimulator entwickelt werden, an dem laparoskopisch-chirurgische Eingriffe oder Eingriffe verwandter Fachrichtungen geübt werden können. VIRTOP steht dabei für VIRTually OPeration training. Die Besonderheit des zu entwickelnden Simulators bestand darin, dass die graphische Darstellung und Simulation der Organe mit einer Echtzeit-Gefühlsübertragung im Sinne einer realistischen Haptik korreliert werden sollte, so dass der Benutzer in 6 Freiheitsgraden die Gewebestruktur ertasten kann. Mit Beendigung des VIRTOP-Projekts konnte noch kein fertiger Simulator erzeugt werden, daher sind weitere Folgeprojekte geplant. Einen Schwerpunkt bildet dabei die Erfassung der Organgeometrie und deren Speicherung als 3D-Modell.

Da Organe ihre Form außerhalb des Körpers verändern, ist es besonders wichtig, dass sie innerhalb des Körpers vermessen werden. Aktuell ist die Vermessung mittels tomographischer Verfahren Stand der Technik. Dem Vorteil der

volumenhaften Erfassung mit diesen Verfahren stehen die Nachteile einer geringen zeitlichen Auflösung und einer Strahlenexposition entgegen. Zudem ist eine zeitgleiche Messung von Krafterwirkungen und der korrespondierenden geometrischen Änderung der Organoberfläche im Sinne einer Kraft-Weg-Messung nicht möglich. Objekte, welche eine Krafterwirkung auf die Organe ausüben, können bei tomographischen Verfahren zu Störungen führen. Neben tomographischen Verfahren eignen sich auch photogrammetrische 3D-Messverfahren zur Erfassung von Organoberflächen. Allerdings müssen die Organe dafür nach aktuellem Stand außerhalb des Körpers untersucht werden. Je nach Messverfahren liefern sie eine hohe zeitliche oder örtliche Auflösung.

2 Material und Methoden

Das Verfahren zur Bestimmung der Oberflächengeometrie, welches in dieser Arbeit vorgestellt werden soll, teilt sich in die vier Phasen auf, wie sie in der Abb. 1 dargestellt sind. In der 1. Phase werden die grundlegenden Schritte der Vorverarbeitung, wie die Kalibrierung, vorgenommen. Sie erfolgt mittels eines Kalibrierfeldes, auf dem Kreismarken angebracht sind. Der relative Abstand der Kreismarken zueinander ist bekannt. Über einen so genannten Bündelblockausgleich erfolgt die Bestimmung der inneren und äußeren Kameraparameter. Das Kalibrierfeld besitzt in Z-Richtung einen Stempel, auf dem ebenfalls Marken angebracht sind. Somit genügt eine Stereobildaufnahme zur Kalibrierung. Die nach der Kalibrierung vorliegende Abweichung zwischen Ist- und Soll-Position der Kreismarken wird als Residuen bezeichnet. Witte beschreibt die Thematik der Kalibrierung ausführlich in seiner Arbeit [2]. Hier finden sich auch weiterführende Informationen zum aktuellen Stand der Technik. Die 2. Phase dient zur Ermittlung von Näherungswerten. Hierzu wird ein Bildpaar aufgenommen und rektifiziert. Zusätzlich werden auch die inneren und äußeren Kameraparameter, wie sie bei Luhmann [3] beschrieben werden, gespeichert und ebenfalls rektifiziert. Aus den rektifizierten Daten wird eine Disparitätenkarte erstellt, deren Werte zur Bestimmung von Tiefeninformationen dienen. Diese Daten werden anhand der Kameraparameter vom Stereonormalfall in den Realfall zurückgerechnet und als Raumpunkte gespeichert. Die so erzeugte Punktwolke ist Ausgangsdatenmenge für die dritte Phase, in der die Punktwolke in ein äquidistantes Punktraster überführt wird. Dafür wird sie mittels Delaunay-Triangulation in ein Oberflächenmodell umgewandelt, welches mit einem äquidistanten Strahlenmuster zum Schnitt gebracht wird. Jeder Schnittpunkt zwischen Oberfläche und Strahl erzeugt einen neuen Raumpunkt. Die Summe aller auf diese Weise ermittelten Raumpunkte ergibt wiederum eine Punktwolke, nun mit äquidistantem Abstand, welche Ausgangsdatenmenge für die Phase 4 ist.

Die 4. Phase beschreibt die Umsetzung des Algorithmus von Michaelis und Albrecht [1] auf dem neuen äquidistanten Punktmuster. Es wird um einen Punkt aus dem Punktmuster eine virtuelle Ebene aus 10 mal 10 Punkten erzeugt. Die Anzahl der Punkte ist frei wählbar, allerdings haben sich Größen zwischen 7 und 20 Punkten je Dimension als geeignet erwiesen. Die aus den Punkten erzeugte

Ebene kann unter anderem über zwei Neigungsparameter und einen Höhenparameter variiert werden. Ziel ist es über eine Optimierung dieser drei Parameter die Ebene maximal an die reale Oberfläche anzunähern. Der Mittelpunkt der Ebene wird durch den 3D-Punkt aus dem äquidistanten Punktmuster beschrieben, er ist auch gleichzeitig der Startwert für die Höhe. Über die ermittelten Kameraparameter werden die Punkte aus dem Objektbereich in die Bildebene projiziert. Dabei liegen die projizierten Punkte meist zwischen mehreren Pixeln. Durch Subpixelinterpolation kann jedem Punkt ein eindeutiger Grauwert zugeordnet werden. Die Gesamtheit all dieser Grauwerte ergibt eine Grauwertmatrix mit entsprechender Größe. Da die Ebene in beide Kamerabilder projiziert wird, ergeben sich zwei Grauwertmatrizen, die mittels normierter mittelwertfreier Kreuzkorrelationsfunktion miteinander verglichen werden. Es wird davon ausgegangen, dass bei maximaler Übereinstimmung die Ebene optimal an die reale Oberfläche angepasst wurde. Um eine maximale Approximation zu erhalten werden die bereits beschriebenen Neigungs- und Höhenparameter variiert. Die Abb. 2 zeigt den vollständigen Kreislauf.

Die Anordnung des Versuchsaufbaus folgt im Wesentlichen dem für die Nahbereichsphotogrammetrie typischen Aufbau, wie sie auch von Luhmann beschrieben wird [3]. Sie besteht aus zwei Kameras zur Aufnahme der Stereobildpaare und einem Beamer als Beleuchtungseinrichtung. Die Optiken der Kameras wurden durch spezielle Endoskope ersetzt, was den Einsatz im menschlichen Körper simulieren soll. Der Abstand zwischen den Endoskopen zueinander sowie der Abstand zwischen Endoskop und Versuchsobjekt wurde stets so gewählt, dass es auf diese Weise auch beim Menschen eingesetzt werden kann. Zur Lösung des Korrespondenzproblems bei einer sehr homogenen Eigentextur der Versuchsobjekte kann mit Hilfe eines Beamers auch ein Muster projiziert werden. Wegen der höheren Auflösung wurden zusätzlich auch einige Dias mit unterschiedlichen Rauschmustern erzeugt und mittels Diaprojektor auf die Objekte projiziert. Betrieben wird das System mit einem handelsüblichem PC.

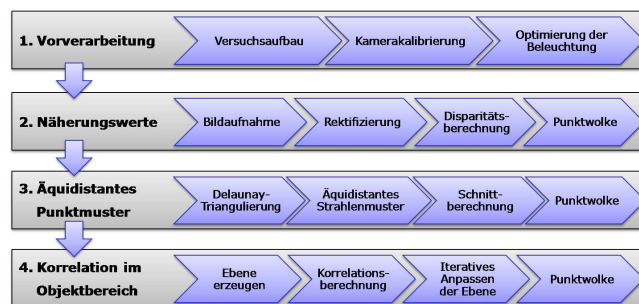


Abb. 1. Das in dieser Arbeit vorgestellte Verfahren in vier Phasen entsprechend ihrer Reihenfolge unterteilt.

3 Ergebnisse

Mit dem vorgestellten Verfahren wurden Organe und andere Probekörper ex vivo, also außerhalb des menschlichen Körpers vermessen. Dabei kann eine Genauigkeit von 0,1 mm bis 1 mm erreicht werden. Durch die Verwendung von Endoskopen ist aber grundsätzlich auch eine in vivo Vermessung möglich. Als Referenzwerte zu den ermittelten Ergebnissen dienen zum einen mit einer Mikrometerschraube real gemessene Werte der Probekörper und zum anderen mit einem hochgenauen mehrschrittigen Verfahren (Phasenkorrelation nach Lilienblum [4]) berechnete Werte. Mit dem Verfahren von Lilienblum ist es ebenfalls möglich, eine Punktwolke für die Versuchsobjekte zu bestimmen. Allerdings ist bei diesem Verfahren die zeitliche Auflösung schlechter.

Probleme entstehen besonders an Organkanten und an starken Zerklüftungen auf der Organoberfläche. Bei einem, mit dem hier vorgestellten Verfahren, vermessenen Rinderherz musste zur Lösung des Korrespondenzproblems ein stochastisches Punktmuster auf die Oberfläche projiziert werden. Untersuchungen haben ergeben, dass die Qualität der Ergebnisse nicht davon abhängig ist, ob das Rauschmuster durch einen Beamer oder Diaprojektor projiziert wird. Beim Beamer ist jedoch darauf zu achten, dass er leicht unscharf zu stellen ist, da es sonst Probleme im Zwischenbereich der Pixel gibt.

Besondere Probleme ergeben sich bei der Vermessung in dunklen, homogenen Bereichen, die das projizierte Muster nur schlecht wiedergeben. Beim Rinderherz, wie auch bei menschlichen Organen, gilt dies im Besonderen für das Muskelgewebe, welches wenig Fett enthält. Hier können erhebliche Probleme bei der Lösung des Korrespondenzproblems auftreten. Abhilfe schafft hier eine stärkere Beleuchtungseinrichtung. Die Abb. 3 zeigt dies für einen besonders dunklen und zerklüfteten Bereich des Rinderherzens.

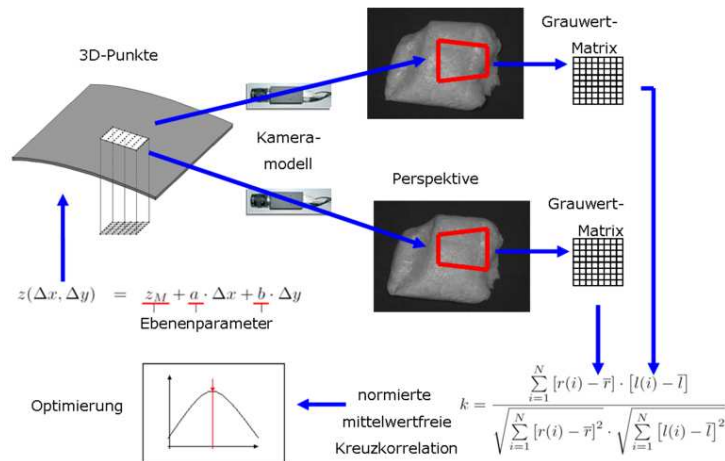
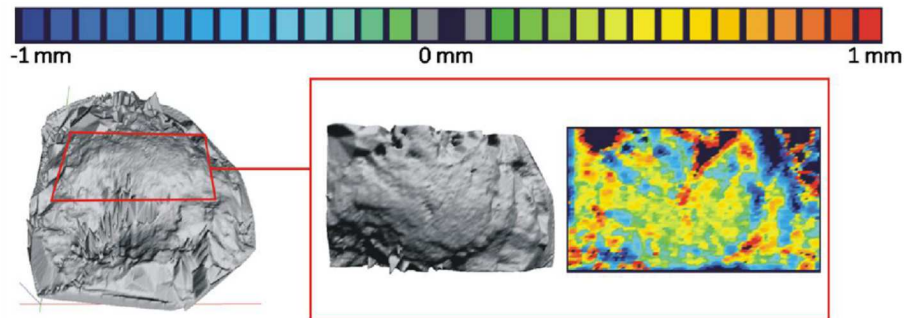


Abb. 2. Flächenkorrelation im Objektbereich nach Michaelis und Albrecht.

Abb. 3. Oberflächenmodell vom Rinderherz mit Falschfarbendarstellung.

4 Diskussion

Für in vivo Anwendungen muss in weiteren Arbeiten geklärt werden, wie das Muster auf das Objekt mittels Endoskop projiziert werden soll. Die Beleuchtungseinrichtung, die in den Endoskopen integriert ist, besteht aus einem Glasfaserbündel. Erste Versuche, durch Eliminierung einzelner Fasern ein Rauschmuster zu projizieren, sind gescheitert. Grund hierfür ist vermutlich die Lichtbeugung am Rand der verbliebenen Fasern, wodurch sich eine geschwächte aber dennoch sehr gleichmäßige Beleuchtung ergibt. Um eine Miniatur der verwendeten Dias vor die Beleuchtung der Endoskope anzubringen, hat deren Lumineszenz nicht ausgereicht. Vermutlich stellt dieses Verfahren jedoch den vielversprechendsten Ansatz dar.

Das Verfahren von Michaelis und Albrecht birgt Schwächen bei stark gebogenen Oberflächen und harten Kanten. Im Extremfall können unerwünschte aber deutlich sichtbare Glättungen an diesen Stellen auftreten. Allerdings besitzen Organe meist eine weiche und eher konvexe Formgebung, was diesen Aspekt relativiert. Probleme existieren weiterhin in allen Bereichen, die nicht von beiden Kameras zeitgleich eingesehen werden können. Abhilfe könnte hier die Verwendung von mehr als nur zwei Kameras schaffen. Allerdings stellt sich in diesem Zusammenhang die Frage, in wie fern der Einsatz in einem Messendoskop und somit im Körper realisiert werden kann.

Literaturverzeichnis

1. Witte M. Photogrammetrische 3D-Vermessung von Organen. Magdeburg: Diplomarbeit, Otto-von-Guericke-Universität Magdeburg; 2009.
2. Albrecht P, Michaelis B. Erhöhung der örtlichen auslösung eines auf flächenkorrelation basierenden 3D-Meßverfahrens. Proc DAGM. 1996; p. 271–80.
3. Lilienblum E, Michaelis B. Optical 3D surface reconstruction by a multi-period phase shift method. J Comput. 2007; p. 73–83.
4. Luhmann T. Nahbereichsphotogrammetrie: Grundlagen, Methoden und Anwendungen. Heidelberg: Herbert Wichmann Verlag; 2000.

Segmentation-Enhanced Registration of Angiography Data

Silvia Born, Daniela I. Wellein, Antje Zöllner, Dirk Bartz

ICCAS/VCM, Universität Leipzig
silvia.born@medizin.uni-leipzig.de

Abstract. For diagnostic and therapeutic purposes, multiple imaging techniques can be used to gain more information on a patient’s anatomy. In maxillofacial and neurosurgery, three-dimensional rotational angiography (3DRA) images can provide a detailed view on blood vessels, whereas computed tomography angiography (CTA) images are common for multiplanar interpretation of bone, soft tissue, and blood vessels. Thus, the registration of 3DRA to CTA allows for a better understanding of the structure and location of blood vessels within a context of bone and tissue, which is essential to surgical interventions. Among other reasons, the lack of mutual information between 3DRA and CTA images makes their registration challenging. In this work, we describe an approach that is based on a segmentation of common structures of the datasets, which enhances the mutual information and hence, the registration result. An evaluation and a comparison to the registration accuracy of unsegmented datasets is presented.

1 Introduction

Three-dimensional rotational angiography (3DRA) provides detailed 3D information, needed to examine a patient’s vasculature from arbitrary view points and to get an understanding of its individual shape. To acquire additional information about surrounding bone and soft tissue, which is essential for surgical intervention planning, further imaging techniques, such as CT angiography (CTA), are used. Here, however, the imaged vasculature is less detailed and these modalities do not allow for an easy phase control (i.e. imaging of only arteries or veins). Accordingly, for the surgeon, a combined exploration of 3DRA and CTA data enables detailed information on blood vessels (from 3DRA) in the context of bone and tissues (from CTA). Here, we focus on the registration of 3DRA to CTA data of the head, which is not straight-forward, due to the different spatial resolutions of the datasets, the fact that the datasets are often taken from different viewing angles, and especially the lack of mutual information in these modalities (Fig. 1 a,b). Therefore, we enhance their mutual information by segmenting structures represented in both datasets (i.e., vessels) and achieve a more accurate registration result when using these segmented gray value images as input.

2 Material and Methods

Mutual Information (MI) methods are considered standard for multimodal rigid registration problems [1] and we therefore based our approach on the MI algorithm introduced in [2]. However, there is little mutual information between 3DRA and CTA [3], which renders MI registration difficult. We enhance the mutual information by segmenting the shared anatomical structures of the two datasets. In this case, the major vessels and bone structures constitute overlap information, because of angiography information in the CTA image and some (low-contrast) representation of bone in 3DRA. In contrast, the inhomogeneous background structures of 3DRA – providing only a low signal-to-noise ratio [4] – decrease the quality of registration results (Fig. 1 a). Preprocessing of the 3DRA dataset is implemented using a median filter at first – to reduce noise while preserving edges – and afterwards apply a region-growing to select bone and vessel structures (Fig. 1 c). For the CTA data, a vesselness filter [5] is used to extract blood vessel information. The final segmentation of vessels and bone is also performed with region-growing (Fig. 1 d). For the selection of the seed points user interaction is required. The thresholds for the other segmentation methods, such as region-growing or vesselness filtering are pre-defined, but can be adjusted by the user if needed.

Our method also allows for an optional initialization step. Here, the spatial correspondence between user-defined landmarks is calculated by minimizing a least-squares distance cost function. Since it is inherently difficult to choose landmarks in two datasets collected at different spatial resolutions and from different viewing angles [6], the segmentation step simplifies the choice of identical landmarks.

A modified version of the *full circle consistency test* as described in [3] is used to determine a quality index e (in millimeter), representing the method's accuracy. Three registrations are performed between three datasets (Fig. 2): 3DRA dataset, original CTA dataset (CTA1), and a dataset (CTA2) obtained by rotating CTA1 by 10° around the X axis and translating it 5 mm in each

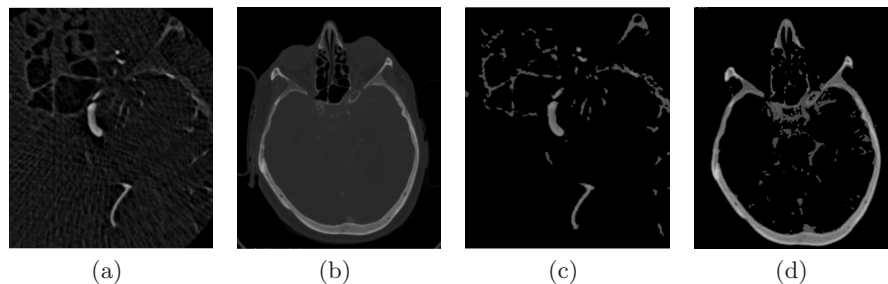


Fig. 1. The modalities 3DRA (a) and CTA (b) differ concerning resolution, field of view and mutual structures. On the right, the respective segmentations of 3DRA (c) and CTA (d) are shown as they are used as registration input.

direction. We let M_{3DRA}^{CTA1} be the transformation of 3DRA dataset to the CTA1 dataset (M_{CTA1}^{CTA2} and M_{CTA2}^{3DRA} named analogously).

For a perfectly accurate registration without any errors, the following will hold:

$$M_{CTA2}^{3DRA} * M_{CTA1}^{CTA2} * M_{3DRA}^{CTA1} = I \quad (1)$$

where I is the identity matrix. If there are any errors, the accumulative error will be equal to

$$E = I - M_{CTA2}^{3DRA} * M_{CTA1}^{CTA2} * M_{3DRA}^{CTA1} \quad (2)$$

From this matrix E , the rotational and translational errors can be extracted. For a better comparability, these errors are transferred into the quality index e converting six error values into a single measure (for details, see [3]). A quality index below 2 mm corresponds to a satisfying registration result, whereas a quality index below the datasets' maximal voxel size is considered excellent.

Four pairs of patient datasets (3DRA and CTA) were available to perform the preprocessing and the described accuracy measurement. With that, the segmentation-enhanced registration was evaluated and compared to the results of the same registration applied to the original (i.e. unsegmented) datasets.

3 Results

After preprocessing, the 3D landmark initialization step is used for a preliminary registration. As this step is based solely on finding the optimal transformation between user-chosen landmarks, it performs as well as the integrity of the chosen landmarks. As can be seen in Figure 3a, the initial registration aligns the images reasonably well, but there is still significant mismatch. The subsequent MI registration finds the optimal transformation. A preliminary visual inspection of the resulting overlay shows a very good registration result (Fig. 3b) promising a high accuracy. Table 1 shows the difference in accuracy of the segmentation-enhanced registration and the registration of the original datasets. In each case, identical landmarks were used for the segmented and original datasets to allow a precise comparison of the results. With a quality index e below or equal to

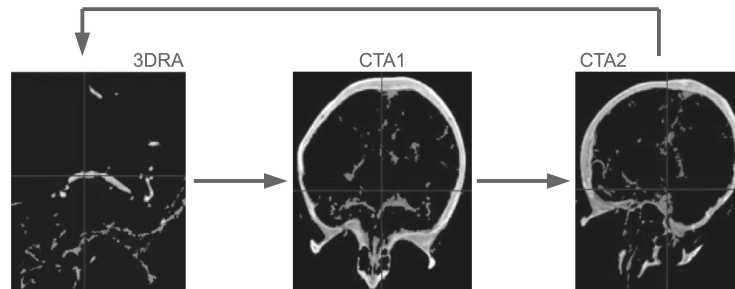


Fig. 2. Registration circle used for evaluation.

Table 1. Registration accuracy in four cases for segmentation-enhanced registration (second column) and the original method (third column). The right-most column depicts the maximal voxel length of the registered datasets.

Case	Quality Index e [mm]		Maximal Voxel Size [mm]
	original	segmented	
1	0.53	0.36	0.7
2	0.92	0.51	0.7
3	1.87	1.29	0.53
4	2.02	0.87	0.53

2 mm, the registration accuracy is satisfying for the original datasets (Section 2). In case 1, a very good result of 0.53 mm, which is below the maximal voxel size of 0.7 mm (right-most column) is achieved. Compared to this, segmenting the mutual structures of the datasets increased the accuracy by at least 32%, in case 4 even by 56%.

4 Discussion

The registration of 3DRA to CTA allows the visualization of vascular details within a context of bone and soft tissue, giving a better understanding of the structure and location of blood vessels. Such a registration is essential for the planning of surgical procedures [7] and has to be fast and highly accurate to meet the needs of the clinical use. Due to the many challenges of registering 3DRA to CTA, an automatic registration can be difficult. In this paper, we presented a successful semi-automatic framework for the registration of volumetric angiography datasets using segmentation and a landmark initialized MI rigid registration algorithm. We showed that enhancing the mutual information by

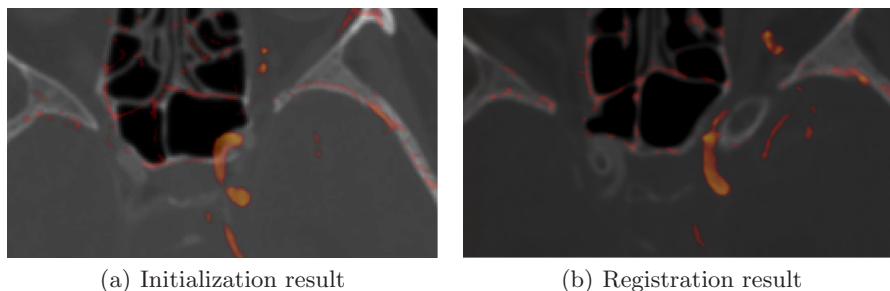


Fig. 3. The landmark initialization step registers the images according to the user-defined landmarks (a). The MI registration refines the registration based on the gray value information of the segmented images (b). (Shown for Case 1; segmented 3DRA data over CTA dataset).

segmenting the datasets' common structures improves registration accuracy. A convenient side-effect is that the segmentation may simplify landmark selection.

Although the preprocessing results in a time overhead compared to using the original datasets directly, the findings are promising. More evaluations concerning the robustness of the method, the influence of the segmentation exactness as well as the number and position of landmarks on the registration results are planned. Furthermore, this approach is not limited to head datasets or CTA data. Thus, an application to heart datasets as well as magnetic resonance angiography data is planned for the near future. Still, further work is needed to reduce the segmentation overhead in order to make this approach competitive to faster methods in the clinical application. For this, we aim at a more automatic segmentation and an automatization of the landmark initialization. In contrast to the quite invasive stereotactic frame registration of [8], we envision a process based on the automatic detection and identification of fiducial markers.

References

1. Hajnal J, Hill D, Hawkes DJ, editors. *Medical Image Registration*. CRC Press; 2001.
2. Mattes D, Haynor DR, Vesselle H, et al. Non-rigid multi-modality image registration. *Med Imag.* 2001; p. 1609–20.
3. Stancanella J, Cerveri P, Colombo F, et al. Development and validation of a CT–3D rotational angiography registration method for AVM radiosurgery. *Med Phys.* 2004;31(6):1363–71.
4. Meijering E, Niessen W, Weickert J, et al. Diffusion-enhanced visualization and quantification of vascular anomalies in three-dimensional rotational angiography: results of an in-vitro evaluation. *Med Image Anal.* 2002;6(3):217–35.
5. Sato Y, Nakajima S, Atsumi H, et al. 3D multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. In: *Proc CVRMed-MRCAS*; 1997. p. 213–22.
6. Sturm B, Powell KA, Stillman AE, et al. Registration of 3D CT angiography and cardiac MR images in coronary artery disease patients. *Int J Cardiovasc Imaging.* 2003;19:281–93.
7. Aoki S, Sasaki Y, Machida T, et al. 3D-CT angiography of cerebral arteriovenous malformations. *Radiat Med.* 1998;16:263–71.
8. Colombo F, Cavedon C, Francescon P, et al. Three-dimensional angiography for radiosurgical treatment planning for arteriovenous malformations. *J Neurosurg.* 2003 3;98:536–43.

Fiber Selection from Diffusion Tensor Data based on Boolean Operators

D. Merhof¹, G. Greiner², M. Buchfelder³, C. Nimsky⁴

¹ Visual Computing, University of Konstanz, Konstanz, Germany

² Computer Graphics Group, University of Erlangen-Nuremberg, Erlangen, Germany

³ Department of Neurosurgery, University of Erlangen-Nuremberg, Erlangen,
Germany

⁴ Department of Neurosurgery, University Hospital Marburg, Marburg, Germany
`dorit.merhof@inf.uni-konstanz.de`

Abstract. Diffusion tensor imaging (DTI) allows investigating white matter structures in vivo which is of particular interest for neurosurgery. A promising approach for the reconstruction of neural pathways are streamline techniques, which are commonly referred to as fiber tracking. However, the resulting visualization of fibers within the whole brain may be complex and difficult to interpret. For this reason, a novel strategy for selecting specific tract systems based on user-defined regions of interest (ROIs) and Boolean operators is presented in this work. The approach provides ultimate flexibility and is an excellent tool for fiber tract selection and planning in neurosurgery.

1 Introduction

Diffusion tensor imaging (DTI) is a magnetic resonance acquisition technique which provides valuable information about the location of white matter tracts within the human brain in vivo. DTI allows quantifying diffusion of water which is anisotropic in tissue with a high degree of directional organization such as in the white matter of the brain. In this way, information about the location of neuronal fibers is provided which is of great interest for neurosurgery.

The reconstruction of white matter tracts from DTI data is commonly solved by tracking algorithms [1, 2] which use streamline techniques known from flow visualization. This tracking strategy provides an intuitive understanding of the spatial relation between fibers and space occupying lesions and is therefore a valuable supplement for neurosurgical planning [3, 4].

Streamlines are traced starting from seed points located in areas of anisotropic diffusion. The resulting fibers represent the white matter tracts within the whole brain, rather than individual tract systems. For this reason, strategies for selecting specific tract systems have been presented. Automated approaches based on clustering [5] do not require any user interaction, but are of limited value for medical application since the fiber selection and clustering is purely based on geometric criteria. For this reason, a more practical approach employs regions of interest (ROIs) defined by the user. In a basic approach [6]

the user is allowed to define multiple box-shaped ROIs, and all fibers crossing the ROIs are displayed. A more advanced approach is based on Boolean selection strategies [7, 8], which are however restricted to two ROIs only.

For planning in neurosurgery, these approaches are not sufficient and more sophisticated selection methods are required. For this reason, an approach based on an arbitrary number of user-defined ROIs with associated Boolean operators is presented in this work, which provides ultimate flexibility and user-interaction.

2 Material and Methods

2.1 Image Data

The DTI datasets used in this work were acquired on a Siemens MR Magnetom Sonata Maestro Class 1.5 Tesla scanner. The specifications of the gradient system were a field strength of up to 40 mT/m (effective 69 mT/m) and a slew rate of up to 200 T/m/s (effective 346 T/m/s) for clinical application. The field of view was 240 mm, resulting in a voxel size of $1.875 \times 1.875 \times 1.9 \text{ mm}^3$. For each of the six diffusion weighted datasets (gradient directions $(\pm 1, 1, 0)$, $(\pm 1, 0, 1)$ and $(0, 1, \pm 1)$) and the reference dataset, sixty slices with no intersection gap and an acquisition matrix of 128×128 pixels were measured. Additionally, a MRI_{T1} sequence ($256 \times 256 \times 160$ voxels) was acquired in each patient which was used for surgical planning.

2.2 Fiber Tracking

Fiber tracking algorithms based on streamline propagation have several steps in common such as seed point selection, fiber propagation and termination strategies [9]. Starting from predefined seed points, streamline integration is used to propagate the fiber until a termination criterion is reached. For this purpose, a threshold based on fractional anisotropy (FA) [1, 2] is commonly applied to stop fiber propagation if the anisotropy of diffusion decreases.

In this work, fiber tracts were computed using fourth order Runge-Kutta integration and trilinear tensor interpolation. A constant tracking step size of 0.5mm was employed. FA was used as termination threshold for fiber propagation, and a length threshold of a minimum of 100 tracking steps was used in order to exclude short fibers from the tracking result.

2.3 Definition of Regions of Interest

In order to extract individual fiber tracts, the user needs to define regions of interest (ROIs) that are used in the fiber selection process. The ROIs can have arbitrary shape and are defined using a voxel-based segmentation tool. The DTI dataset that is measured without diffusion gradient is used as anatomical scan and loaded into the segmentation tool. The ROI is drawn manually on axial, sagittal or coronal sections based on anatomical knowledge. In this way, also non-trivial ROI shapes can be defined, e.g. in the vicinity of a tumor. Finally, the ROI is imported into the fiber tracking tool for further processing.

2.4 Boolean Operators associated to Regions of Interest

Boolean logic has many applications in electronics, computer hardware and software, and is the basis of all modern digital electronics. Boolean operators are used in order to define conditions on sets in order to retrieve the qualifying elements. The basic Boolean operators are logical OR and logical AND, which perform the union and intersection of sets, and logical NOT which is a unary operator and forms the complement.

In order to further control the fiber selection process, each ROI has an associated Boolean operator that defines a rule for fiber tract selection for this ROI. Basically, the following three options exist:

- *AND*: A ROI with associated Boolean operator AND requires any fibers that are retained to cross this ROI. Multiple AND ROIs result in a fiber bundle that crosses all AND ROIs.
- *OR*: A ROI with associated Boolean operator OR requires any fibers that are retained to either cross this ROI or any other OR ROI. If only one OR ROI is defined, the effect is the same as if the OR ROI was an AND ROI.
- *NOT*: In some cases, it is desirable to exclude certain fibers from the extracted fiber tract, e.g. fibers that connect to the other hemisphere. For this purpose, a NOT ROI may be defined, which excludes any fibers that cross the ROI.

By combining ROIs with different associated Boolean operators, utmost flexibility for tract selection is provided. In Fig. 1, the user interface for fiber tracking is shown, the associated Boolean operator can be toggled for each ROI. The approach has been integrated into the software framework MedAlyVis (Medical Analysis and Visualization).

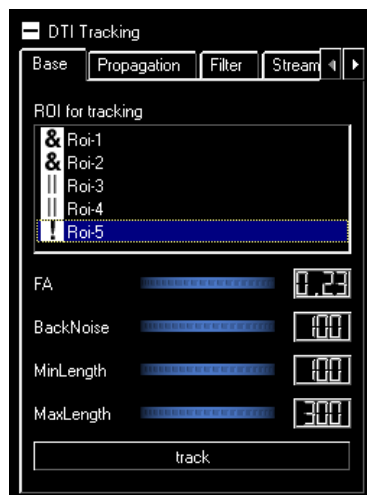


Fig. 1. Fiber tracking tool. The Boolean operator can be toggled for each ROI. Available options: AND, OR, NOT.

3 Results

In Fig. 2, different combinations of ROIs with associated Boolean operators are shown for a brain tumor patient. If no ROI is defined, the fibers resulting from fiber tracking within the whole brain are displayed (upper left). Combination of two AND ROIs allows extracting the pyramidal tract (upper right). A combination of an AND ROI and two OR ROIs allows displaying the fraction of fibers that expand from the AND ROI to either of the OR ROIs (lower left), and a NOT ROI is used in order to exclude undesirable fibers that belong to a different tract system (lower right).

The performance of the tract selection method depends on the number of voxels within the dataset, the number of fibers, the size of the ROIs and the individual position of the ROIs. For a whole brain tracking comprising 300.000 fibers, the fiber tracking and fiber tract selection using up to four ROIs with different associated Boolean operators took between 20 and 50 seconds on a PC equipped with a Intel Core i7 CPU (3.0 GHz) and 3 GB RAM.

4 Discussion

Fiber selection techniques are of major importance in order to make fiber tracking results accessible for surgery planning and intra-operative visualization. The presented approach overcomes the limitations of automated clustering strategies [5], manual selection techniques [6], and Boolean selection strategies [7, 8] that were presented previously.

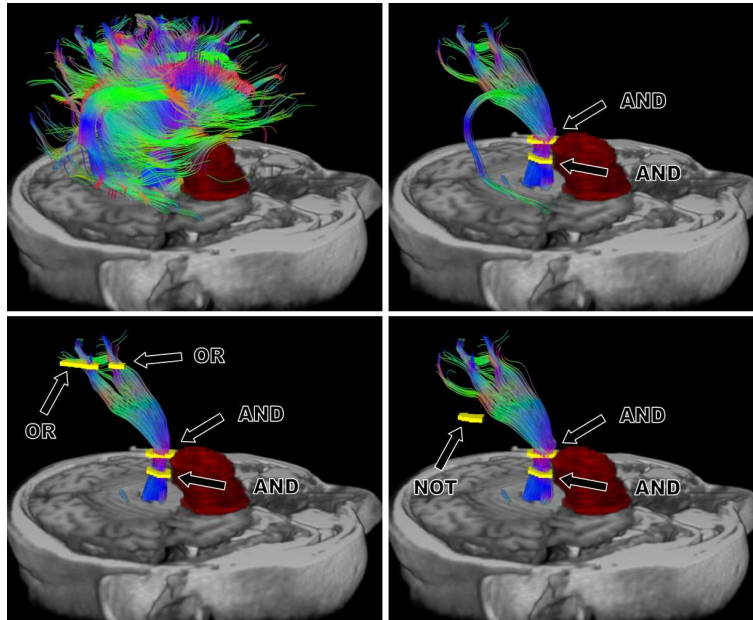
Compared to the timings reported in [6], the approach is slower since a tracking of the whole brain is always performed in the first instance. Options to speed up the approach would be to compute the fibers within the whole brain only once, and to use an octree in order to speed up the fiber selection [6].

Nevertheless, the possibility of dedicated tract extraction for surgery planning was greatly appreciated by the neurosurgeons, and the computing times were still considered as acceptable. Especially the option to exclude perturbing fibers using NOT ROIs proved to be very desirable.

5 Conclusion

In this work, a novel strategy for selecting specific tract systems based on user-defined ROIs and Boolean operators is presented, which is applicable to an arbitrary number of ROIs. The approach overcomes the limitations of current techniques and offers ultimate flexibility for tract extraction. Overall, the presented technique makes it possible to extract fiber tracts in a very patient-specific way, which is desirable for surgical planning and intra-operative visualization in neurosurgery.

Fig. 2. Fiber tracts extracted by using a combination of ROIs with different associated Boolean operators. Upper Right: Fiber tracking result without ROIs. Upper Right: Usage of AND ROIs. Lower Left: Effect of using OR ROIs. Lower Right: Undesirable fibers are excluded using a NOT ROI.



References

1. Basser PJ, Pajevic S, Pierpaoli C, et al. In vivo fiber tractography using DT-MRI data. *Magn Reson Imaging*. 2000;44(4):625–32.
2. Mori S, Crain BJ, Chacko VP, et al. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann Neurol*. 2001;45(2):265–9.
3. Clark C, Barrick T, Murphy M, et al. White matter fiber tracking in patients with spaceoccupying lesions of the brain: a new technique for neurosurgical planning? *Neuroimage*. 2003;20(3):1601–8.
4. Nimsy C, Ganslandt O, Hastreiter P, et al. Intraoperative diffusion-tensor MR imaging: shifting of white matter tracts during neurosurgical procedures: initial experience. *Radiol*. 2005;234(1):218–25.
5. Moberts B, Vilanova A, van Wijk JJ. Evaluation of fiber clustering methods for diffusion tensor imaging. *Proc IEEE Vis*. 2005; p. 65–72.
6. Blaas J, Botha CP, Peters B, et al. Fast and reproducible fiber bundle selection in DTI visualization. *Proc IEEE Vis*. 2005; p. 59–64.
7. Jiang H, van Zijl P, Kim J, et al. DtiStudio: Resource program for diffusion tensor computation and fiber bundle tracking. *Comput Methods Programs Biomed*. 2006;81(2):106–16.
8. Chao Y, Chen J, Cho K, et al. A multiple streamline approach to high angular resolution diffusion tractography. *Med Eng Phys*. 2008;30(8):989–96.
9. Mori S, van Zijl PCM. Fiber tracking: principles and strategies: a technical review. *NMR Biomed*. 2002;15(7-8):468–80.

Evaluation of a Time-of-Flight-based Respiratory Motion Management System

Christian Ulrich, Christian Schaller, Jochen Penne, Joachim Hornegger

Pattern Recognition Lab, Friedrich Alexander University Erlangen-Nuremberg
`christian.schaller@informatik.uni-erlangen.de`

Abstract. In this paper a novel anatomic-like phantom, to simulate human respiratory motion is presented. The phantom is capable to simulate thorax and abdomen motion for surface based respiratory gating systems. This phantom is used to evaluate a system based on time-of-flight (TOF) technology, to detect respiratory motion. It could be shown that the correlation between the reference motion, performed by the phantom and the measured data of the TOF system is 0.65 for a breathing amplitude of 1.5 mm and above 0.80 for amplitudes bigger than 5 mm. Furthermore, the system is capable to detect a breathing frequency of at least 25 respiratory cycles per minute.

1 Introduction

In radiotherapy the handling of moving targets is getting more and more important [1]. Moving targets are e.g. tumors which are moving inside the human body due to respiratory motion. These targets can be irradiated efficiently by using respiratory gating. Thereby the tumor is irradiated in a predefined respiratory phase and position. One method for respiratory gating is to acquire an external surrogate respiratory signal. There are several methods to acquire such a signal, e.g., the AZ-733 V by Anzai Medical [2], GateRT by VisionRT [3], the Varian RPM system [4] and a time-of-flight (TOF) sensor-based system [5]. Within this paper an anatomic-like phantom is introduced to evaluate external gating methods by providing reference breathing signals for thorax and abdomen. We are using the phantom to evaluate the TOF-based approach, proposed by Schaller et al. [4] and Müller et al. [6].

2 Materials and Methods

In the following we introduce a phantom to simulate human respiratory motion. It can be used to evaluate surface based respiratory gating systems. Such systems have special requirements on phantoms. The shape of the phantom should be human-like and also the motion of the phantom should correspond to human respiratory motion. Existing phantoms for example the Quasar phantom from Modus Medical Devices Inc., do not meet these requirements as they do not provide the correct geometry and motion. These phantoms are also not able

to simulate various breathing amplitudes and velocities. Due to the human-like surface, the novel phantom can be used for simulating patient positioning processes as well [7].

We manufactured a plaster cast of a male person and added mechanical support to its chassis. This mechanical support is used to simulate various human respiratory motion. An 8-bit AVR Atmega32 microcontroller (μC) with a clock frequency of 16 MHz and 32 kByte flash memory was used. The μC manages the communication with a PC via USB and controls the movement of the phantom's abdomen and thorax. The abdominal motion is generated by a DC-motor with a rotary encoder as control feedback, the thorax is lifted and lowered with three servomotors (Fig. 1). A plaster cast was used, because the intended use of this phantom is to evaluate surface based systems. The intra-abdominal motion of the organs was simplified as a lift and lowering of the abdominal surface. The rotatory movement of the ribs, according to the spine, resulting in a 2D movement of the thorax, is simulated as a rotatory movement by the servo-motors. Using these components various respiratory motions can be simulated. The corresponding reference signals can be used as ground truth for correlation computations and evaluation.

The reference signal is a cosine-shaped signal which simulates human respiratory motion. The thorax performs a rotational movement, whereas the abdomen a linear up and down motion. The phantom can be parameterized using a software tool. Various frequencies and amplitudes for both, thorax and abdomen can be set. It is also possible to simulate the connection of both respiratory systems using a couple factor. In general, the system is also capable to simulate irregular motion like coughing.

3 Results

We used a standard CamCube TOF camera from PMD Technologies GmbH and observed the phantom movements within a distance of about 1 m. The test setup

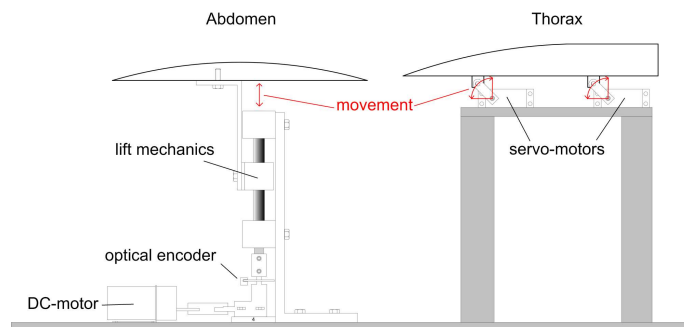


Fig. 1. Schematic overview of the mechanical motion of the thorax and the abdomen (sideview).

is illustrated in figure 2. The TOF camera provides a 3-D representation of the shape and motion of the phantom. Using the method introduced by Schaller et al. [4] two independent respiratory signals for thorax and abdomen can be acquired. The phantom is used to evaluate the capabilities of the TOF-based system.

To investigate different respiratory signals the phantom breathing frequency was varied between $3 - 25 \text{ min}^{-1}$, the phantom abdomen and thorax amplitude from $1.3 - 18 \text{ mm}$. The resting respiratory frequency of an adult is $12 - 15 \text{ min}^{-1}$, during stress, e.g. during radiotherapy the respiratory frequency can increase. On the opposite a sedated patient can have a reduced respiratory activity.

After a gain-offset correction of the measured and reference signal the correlation factor according to Pearson was computed. For synchronization purposes, the TOF signal was shifted two timestamps to match the phantom signal. This phase shift is the result of the non-deterministic thread handling in the current operating system. To acquire the two respiratory signals the USB and the serial port of the PC have to be polled simultaneously.

Several measurements were performed. The correlation factors of the TOF and phantom signal varied from 0.36 up to 0.99 depending on amplitude and frequency. Figure 3(a) illustrates a respiratory signal with a high correlation, whereas figure 3(b) shows a lower correlation.

4 Discussion

Smaller correlation values, having small peak-to-peak amplitudes, can be explained by the noise of the TOF signal in z -direction (Fig. 4(a)). No preprocessing to smooth the TOF data was applied. Therefore, the z -component is noisy. It can be reduced with a mean or bilateral filter, which can increase the correlation and therewith the signal equality of the real respiratory motion and the measured motion. An adequate correlation without filtering can be achieved with a peak-to-peak amplitude of 5 mm and above.

The second parameter of respiratory motion is the frequency. Figure 4(b) shows that the correlation factor can be considered independent of the respira-

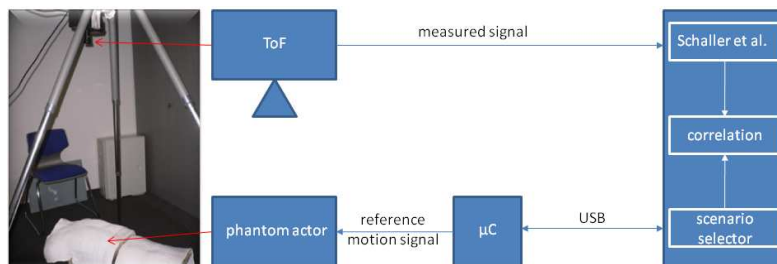


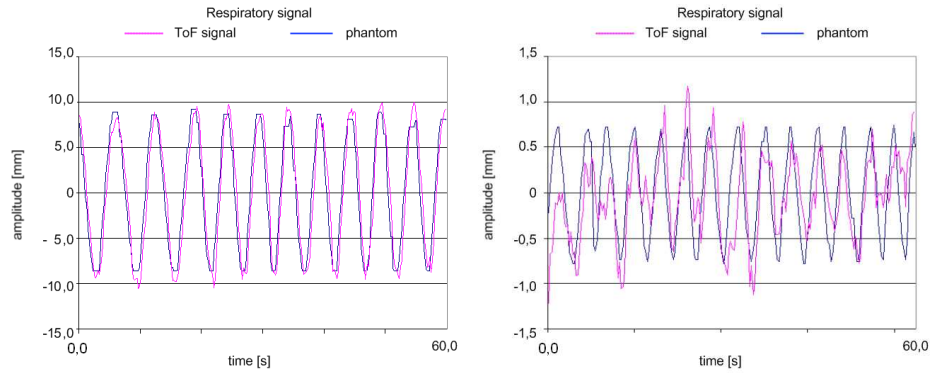
Fig. 2. Schematic overview of the test setup for the evaluation of the respiratory gating system. The system can be used to evaluate surface-based motion detection systems.

tory frequency in the range of $3 - 25 \text{ min}^{-1}$. Concluding, the TOF camera can acquire a stable respiratory signal independent from the respiratory frequency with at least 25 breathings per minute.

Figure 3(b) shows the limitations of respiratory gating with a TOF camera. Here a breathing amplitude of 1.5 mm peak-to-peak with a respiratory frequency of 14 min^{-1} was simulated. Thereby a correlation factor of 0.65 was achieved.

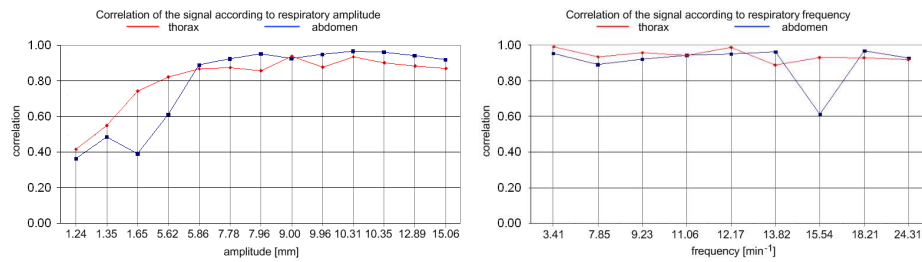
The TOF system returned a stable respiratory frequency and a well fitting amplitude signal. With suitable signal post-processing and filtering of the raw data the signal similarity and thereby the correlation can be furthermore increased.

We introduced an anatomic-like breathing phantom which can simulate custom amplitudes and frequencies to generate reference breathing motion signals. The phantom can be used to evaluate surface based respiratory gating or positioning systems accurately. It provides ground truth signals in real-time. As a case study, we evaluated a TOF based respiratory motion detection system and could show in a quantitative manner that it provides reliable data.



(a) Respiratory frequency of 11 min^{-1} , amplitude of 17.8 mm, correlation factor of 0.94. (b) Respiratory frequency of 14 min^{-1} , amplitude of 1.5 mm, correlation factor of 0.65.

Fig. 3. Evaluation of the thorax movement in ToF camera direction.



(a) Amplitudes (b) Frequency

Fig. 4. Correlation factor of the ToF signal and the phantom signal.

References

1. Keall PJ, Mageras GS, Balter JM, et al. The management of respiratory motion in radiation oncology report of AAPM Task Group 76. *Med Phys.* 2006;33(10):3874–900.
2. Dietrich L, Jetter S, Tücking T, et al. Linac-integrated 4D cone beam CT: first experimental results. *Phys Med Biol.* 2006;51(11):2939–52.
3. Tarte S, McClelland J, Hughes S, et al. A non-contact method for the acquisition of breathing signals that enable distinction between abdominal and thoracic breathing. *Radiother Oncol.* 2006;81(1):S209.
4. Ruan D, Fessler J, Balter J. Mean position tracking of respiratory motion. *Med Phys.* 2008;35(2):782–92.
5. Schaller C, Penne J, Hornegger J. Time-of-flight sensor for respiratory motion gating. *Med Phys.* 2008;35(7):3090–3.
6. Müller K, Schaller C, Penne J, et al. Surface-Based respiratory motion classification and verification. In: *BVM; 2009.* p. 257–61.
7. Placht S, Schaller C, Adelt A, et al. Improvement and evaluation of a time-of-flight-based patient positioning system. In: *BVM2010; 2010.* p. to appear.

Markerbasiertes Online Kalibrierverfahren für die CT-Rekonstruktion

Stephanie Simbt¹, Frank Dennerlein², Jan Boese²

¹Fachbereich EMW, Hochschule Anhalt (FH)

²Siemens AG, Healthcare Sector, Forchheim, Germany

`stephanie.simbt@freenet.de`

Kurzfassung. Das etablierte Verfahren der Offline-Geometriekalibrierung ist bei mechanisch sehr flexiblen Röntgensystemen oftmals problematisch. Durch unregelmäßige, geometrische Abweichungen der 3D-Aufnahmebahn bezüglich der Aufnahmebahn der Offline-Kalibrierung kommt es zu einer Verringerung der Bildqualität der CT-Aufnahmen. Um die Bildqualität auf diesen Systemen zu verbessern wird vorgeschlagen eine zusätzliche Aktualisierung der Geometrieinformation in jeder einzelnen Patientenaufnahme durchzuführen. Das hier implementierte Verfahren benötigt wenige kugelförmige Marker, deren detektierte Projektionen in den Röntgenbildern mit den jeweiligen Röntgenquellenpositionen Strahlen definieren, deren Quasi-Schnittpunkte die 3D Positionen der Marker im Raum ergeben. Es wird vorgeschlagen, die Aufnahmegeometrie so zu korrigieren, dass sich eine Rückprojektion der 3D-Markerpositionen möglichst genau mit den im Bild gefundenen Markerpunkten überdeckt. Tests an simulierten und realen Datensätzen zeigen, dass die Projektionsbildabweichung um min. 79 % reduziert werden kann.

1 Einleitung

Für die Gefilterte Rückprojektion bei der CT-Rekonstruktion werden Informationen über Geometrie und Position des Röntgensystems während der Aufnahmebewegung bezüglich des zu untersuchenden Objektes benötigt. In Standardsystemen ist es meist ausreichend diese Informationen in größeren Zeitabständen mithilfe einer Offline-Kalibrierung des Röntgensystems zu erzeugen. Bei mechanisch sehr flexiblen Systemen, die ebenfalls eine 3D-CT-Bildgebungsfunktion ermöglichen, kann sich die Geometrie der 3D-Aufnahmebahn des Systems aber innerhalb dieses Zeitraums, sogar von Aufnahme zu Aufnahme, verändern. Tritt dieser Fall ein, stimmen die Offline-Kalibrierinformationen nicht mehr mit den tatsächlichen Geometrieparametern überein, was zu Fehlern in der Rückprojektion und damit zu einer schlechteren Bildqualität der berechneten Schnittbilder [1] führt. Ein möglicher Lösungsansatz dazu ist es, in einer Online-Kalibrierung bei jeder Patientenaufnahme die Geometrieparameterabweichungen zur Offline-Kalibrierung zu detektieren und auszugleichen. Eine solche Online-Korrektur der Geometrieinformation geschieht oft mittels röntgenpositiver Marker im Blickfeld

des Röntgensystems. Um eine Verdeckung von Patientenstrukturen durch Metallartefakte zu vermeiden, wird für die Online-Kalibrierung im Folgenden eine Methode untersucht, welche die Verwendung weniger, geometrisch einfacher Marker ermöglicht. Bei dem markerbasierten Verfahren zur Bewegungskompensation des Patientenkopfes nach Jacobson und Stayman [2] sind z.B. mehrere Kugelmarker in einen flexiblen Ring eingefasst, der am Patientenkopf befestigt wird. Zur Schätzung der Patientenbewegung während der CT-Aufnahme werden Rotationsmatrix und Translationsvektor jeder aufgenommenen Projektion anhand der Marker im Bild bestimmt. In Anlehnung an diesen Bewegungskorrekturansatz wurde hier ein Verfahren zur Optimierung einzelner, systemspezifischer Parameter entwickelt, in denen sich die Geometrieabweichungen am stärksten manifestieren (durch gesondertes Verfahren ermittelt). Dieses Vorgehen schränkt den Freiheitsgrad des Korrekturproblems ein, sodass nur eine sehr geringe Anzahl an Markern benötigt wird.

2 Material und Methoden

Das entwickelte Kalibrierverfahren verwendet mehrerer Kugelmarker, soll jedoch am Beispiel eines einzelnen Markers erläutert werden. Gegeben sind die (mittels Standardverfahren der Bildverarbeitung) detektierten Markermittelpunkte in den Projektionsbildern, sowie die potentiell ungenauen Aufnahmegeometrieinformationen in Form von Projektionsmatrizen in homogener Schreibweise (3×4 -Matrizen) [3]. Verbindet man in jeder Projektion die 3D-Koordinate des detektierten Markerpunktes x^d mit der zugehörigen Quellposition a durch eine Gerade, ergibt sich bei idealer Geometrieinformation ein gemeinsamer Schnittpunkt aller Geraden bzw. Strahlen, der die Position des Markers W im Field-of-View widerspiegelt (Abb. 2, links). Bei verfälschten Geometrieinformationen, die durch Benutzung von Offline-Kalibrierdaten bei schlecht reproduzierbaren Aufnahmebahnen entstehen, gibt es jedoch im Allgemeinen keinen eindeutigen Schnittpunkt. Stattdessen kann durch eine Mittelwertsberechnung unter Verwendung der kürzesten Verbindungen zwischen den generierten Strahlen im Raum eine Mittlere Markerposition W' ermittelt werden, welche der realen Position des Markers im Field-of-View sehr nahe kommt (Abb. 2, rechts). Im Allgemeinen kommt es nach der Projektion von W' auf die Detektorfläche mittels der

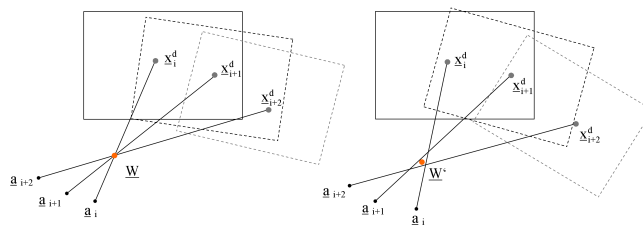


Abb. 1. Strahlenverlauf bei idealer (links) und verfälschter Geometrieinformation (rechts).

vorliegenden Projektionsmatrix zu einer Abweichung zwischen der projizierten Mittleren Markerposition \widetilde{W}' und den detektierten Markerpunkten der Detektorebene \underline{x}^d . Diese Abweichung kann mithilfe der euklidischen Distanz für jede Projektion n angegeben werden. In dem hier vorgeschlagenen Verfahren werden nun zunächst diejenigen Geometrieparameter des Systems bestimmt, in denen sich die Geometrieabweichungen zur Offline-Kalibrierung am deutlichsten manifestieren. Für das hier untersuchte System hat sich die Korrektur zweier Projektionsmatrix-Elemente als geeignet herausgestellt. Ziel der Optimierung ist anschließend das Finden von Korrekturwerten zu diesen Elementen in jeder Projektionsmatrix, so dass die Abstände zwischen Vorwärtsprojektion der Mittleren Markerpositionen \widetilde{W}' und den zugehörigen detektierten Markerpunkten \underline{x}^d minimiert werden (Abb. 2). Für die Korrektur der Geometrieinformation wird zu den c Projektionsmatrixelementen $p = (p_1, p_2, \dots, p_c)$ (hier $c = 2$) ein $\delta = (\delta_1, \delta_2, \dots, \delta_c)$ hinzu addiert. Dieses δ wird mithilfe der euklidischen Distanz zwischen der in der aktuellen Korrekturschätzung projizierten Markerposition

$$\widetilde{W}' = (w_1 \ w_2 \ w_3)^T \quad \text{mit} \quad u_w = \frac{w_1}{w_3} \quad \text{und} \quad v_w = \frac{w_2}{w_3} \quad (1)$$

und der im Projektionsbild detektierten Markerposition u_s und v_s bestimmt. Der Mittelwert dieser euklidischen Distanzen über alle K detektierten Marker in einer Projektion liefert somit das Optimierungskriterium

$$\Psi(\delta) = \frac{\sum_{k=1}^K \sqrt{(u_s^{\{k\}} - u_w^{\{k\}})^2 + (v_s^{\{k\}} - v_w^{\{k\}})^2}}{K} \quad (2)$$

Die Koordinaten u_w und v_w sind von der Projektionsgeometrie und damit von δ abhängig. Wird mithilfe des iterativen Nelder-Mead-Optimierungsverfahrens eine festgelegte Genauigkeit des gefundenen Minimums erreicht, gelten die dazu verwendeten Werte für δ als geeignete Korrekturwerte für die jeweiligen Parameter. Diese Optimierung wird für jede Projektion wiederholt, sodass jede einzelne, mittels Offline-Kalibrierung gewonnene Projektionsmatrix an die aktuell gültige Aufnahmegeometrie angepasst wird. Dies ermöglicht eine genauere Durchführung des Rückprojektionsschrittes im Rekonstruktionsverfahren und führt somit

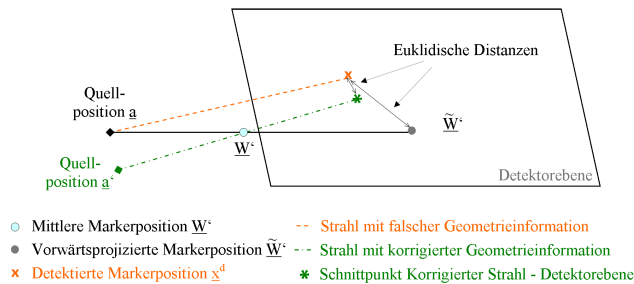


Abb. 2. Prinzip der Geometriekorrektur.

zu einer verbesserten Bildqualität in den rekonstruierten Schichtbildern. Theoretisch ist für die Korrektur der zwei Matrixelemente nur ein einzelner Marker notwendig. Praktisch werden allerdings mehrere Marker verwendet, um eine eindeutige Lösung [3] und ein robusteres Ergebnis zu erhalten. Der oben beschriebene Online-Kalibrieralgorithmus wurde in MATLAB implementiert und u.a. an simulierten Röntgenprojektionen eines virtuellen Kopfmodells (so genanntes Simulationsmodell), an realen Aufnahmen eines ConeBeam-Phantoms und an einer 3D-Röntgenprojektion eines menschlichen Femurs evaluiert. Bei allen Datensätzen sind die optimalen Informationen über die Systemgeometrie bekannt (bei den realen Datensätzen ist diese Information mit einer parallel zur Datenaufnahme durchgeführten Kalibrierung erzeugt worden). Nachträglich wurden verschiedene Systemparameter mittels Zufallswerten gestört und mithilfe des Algorithmus wieder korrigiert, sodass ein direkter Vergleich zwischen optimalem, verfälschtem und korrigiertem Ergebnis in Tab. 1 ermöglicht wird.

3 Ergebnisse

Bei der Aufnahme der Testdatensätze wurden je 5–7 Kugelmarker individuell an der Oberfläche der Teststruktur angebracht (bzw. in dem Simulationsmodell hinzu modelliert). Für die 3D-Rekonstruktion kam das Verfahren nach [4] zum Einsatz, welches eine 3D-Erweiterung des 2D-Rekonstruktionsansatzes nach [5] darstellt. Im oberen Teil von Abb. 2 wird deutlich, dass durch Verfälschung der Geometrieinformationen (mittleres Bild) starke Linienartefakte im Schichtbild des Simulationsmodells auftreten und weder scharfen Kanten, noch feine Strukturen erkennbar sind. Im Bild der korrigierten Geometrieinformation (rechts)

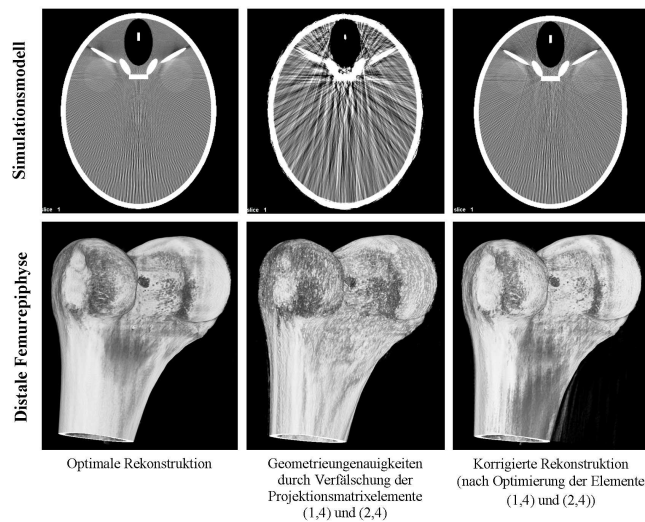


Abb. 3. Auswirkungen der Geometrieinformation auf die Bildqualität.

Tabelle 1. Bildabweichung als Versuchsergebnis im Vergleich.

	Verfälschte Geometrie- information	Korrigierte Geometrie- information	Verbesserung durch Korrektur
Simulationsmodell	5,7763 Pixel	0,4152 Pixel	92,8 %
ConeBeamPhantom	2,6290 Pixel	0,4868 Pixel	81,5 %
Femur	3,1016 Pixel	0,6270 Pixel	79,8 %

dagegen ist nur ein geringer Unterschied zur optimalen Geometrieinformation (links) sichtbar. Auch in der 3D-Darstellung der Femurepiphyse wird sichtbar, dass die Störung der Geometrieinformation die Oberfläche des Knochens stark verfälscht und nach der Parameterkorrektur wesentlich glatter aussieht. Anhand des ConeBeam-Phantoms lässt sich nach Störung der Projektionsmatrizen eine Auflösung von 10 LP/cm feststellen. Nach Korrektur der Geometrieinformation wird dagegen die fast optimale Auflösung von 16–18 LP/cm erreicht.

4 Diskussion

Es ist gelungen ein effizientes, markerbasiertes Online-Geometriekalibrierverfahren für CT-Röntgensysteme zu entwickeln, welches die Verringerung der mittleren Projektionsbildabweichung mittels Korrektur einiger weniger Parameter bewirkt. Die vorgestellte Methode schränkt zuerst den zu korrigierenden Parameterraum auf die bildwirksamen Parameter ein, so dass in der Regel weniger Marker benötigt werden als in dem Verfahren von [2]. Anhand der Verbesserung der Bildabweichung um min. 79 % und dem Vergleich der rekonstruierten Bild-
daten bei allen durchgeführten Experimenten lässt sich von der Wirksamkeit des entwickelten Verfahrens ausgehen. Es wurde somit ein effizientes Verfahren implementiert, mit dem eine hochqualitative 3D-Bildgebung selbst bei Systemen möglich wird, bei denen die Reproduzierbarkeitsannahme nicht gültig ist.

Literaturverzeichnis

1. Kyriakou Y, Lapp RM, Hillebrand L, et al. Image-based online correction of misalignment artifacts in cone-beam CT. *Medical Imaging 2009: Physics of Medical Imaging Proceedings of the SPIE*. 2009;7258:1–10.
2. Jacobson MW, Stayman JW. Compensating for head motion in slowly-rotating cone beam CT systems with optimization transfer based motion estimation. *IEEE Nucl Sci Symp Conf Rec*. 2008; p. 5240–45.
3. Hartley RI, Zisserman A. *Multiple View Geometry in Computer Vision*. Cambridge: Cambridge University Press; 2003.
4. Dennerlein F, Kunze H, Boese J. Cone-beam reconstruction from a variable-radius, planar source trajectory. *Proc IEEE, Med Imaging Con*. 2009 (to be published).
5. Noo F, Defrise M, Clackdoyle R, et al. Image reconstruction from fan-beam projections on less than a short scan. *Phys Med Biol*. 2007;47:2525–46.

2D/3D-Registration of Cerebral DSA Data

Clemens M. Hentschke, Klaus D. Tönnies

Department of Simulation and Graphics, University of Magdeburg
`cmh@isg.cs.uni-magdeburg.de`

Abstract. We present a method to automatically register cerebral 2D digital subtraction angiography sequences (2D-DSA) to 3D rotational X-ray angiography (3D-RA) volumes. The application is the validation of blood flow simulations for aneurysm treatment. We compute projection images of 3D-RA data sets and compare them to the 2D-DSA images. A three-step strategy is presented to calculate the registration where no user interaction is needed. We tested the accuracy of our method on studies of five patients by manually setting fiducial markers and measured a root mean square projection error of $1.78 \text{ mm} \pm 0.45 \text{ mm}$.

1 Introduction

Simulation methods have recently been proposed to support treatment of intracranial aneurysms [1]. Patient-specific simulation of blood flow provides information about the effect of inserted implants such as stents to the blood flow.

Currently, no imaging modality exists that includes spatial and time-resolved blood flow information with high resolution. Time-dependent 2D-DSA displays the blood flow in a projection image. Static 3D-RA volumes incorporate the morphology of vessels. By fusing both modalities, blood flow and spatial information are available. We describe a method to achieve such a 2D/3D-registration.

Although 2D/3D-registration methods exist for quite some time (e.g. [2]), they require appropriate initialization for being successful. Such interaction is not acceptable for use in practice. Hence, our goal is to provide a fully automatic registration algorithm. The lack of user interaction should not affect the precision of our method with regard to similar approaches like [3] and [4].

The main challenges are to cope with the low information content in angiographies and the fact that acquisition geometry is known only approximately.

2D/3D-registration requires a projection of the 3D volume to the 2D image. X-ray based modalities like X-ray fluoroscopy were successfully registered to CT data sets [2]. An intensity-based approach was published that registers 2D-DSA to magnetic resonance angiography volumes [3] and later extended to register also 3D-RA volumes [4]. Recently, a feature-based deformable 2D/3D-registration for vascular models was presented [5].

The disadvantage in many algorithms is the need for user interaction. Several approaches rely on setting an initial position of the 2D image with respect to the 3D volume. In case of feature-based methods, a segmentation is commonly

needed. Our registration requires neither a preprocessing step, nor user intervention. Also, in contrast to other algorithms only the approximate projection geometry has to be known.

2 Material and Methods

We achieve automatic registration by an iterative scheme that finds subsets of correct registration parameters in three steps. We will first describe the registration parameters before we continue with a description of the process.

The first input data set consists of a sequence of n 2D-DSA images that shows the blood propagation. The second input modality is a static 3D-RA volume. We assume that there are no major changes in the cerebral vessel system and a known orientation of both data sets.

The image intensities in both modalities correlate with the measured contrast agent (CA). As a 3D-RA data set shows the distribution of CA in the whole cerebral vessel system, we simulate this behavior by integrating the intensity values in the 2D-DSA sequences over time. The result is denoted as a 2D-MAX image. A projection is computed by raycasting through the 3D volume.

We use the maximum intensity projection (MIP) technique to produce 2D projection images and employ two transfer functions to visualize different vessel aspects. The first function aims to simulate a 2D-MAX image by visualizing large and medium vessels predominantly (MIP₁, Fig. 1(a)). The second function is used for displaying also small vessels and potential noise (MIP₂, Fig. 1(b)).

Registration requires to find intrinsic and extrinsic parameters. Intrinsic parameters can be assumed to be known as they can be estimated from image properties and meta information [4]. Extrinsic parameters describe the rigid transformation parameters and a projection with respect to a world coordinate system. Six extrinsic parameters ($\phi, \theta, t_x, t_y, s, \alpha$) have to be found. ϕ and θ are the projection parameters (Fig. 1(c)). t_x and t_y are the 2D translation parameters, s and α are the 2D scaling and the in-plane rotation, respectively.

Automatic registration is carried out in three steps with 2D-MAX image and projection image as input. Each step focuses on a subset of registration

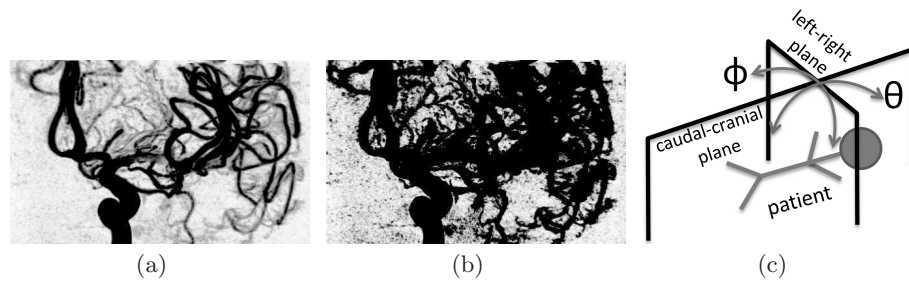


Fig. 1. Example of a projection image created with MIP₁ (a) and MIP₂ (b); scheme of the projection parameters ϕ and θ that define the position of the C-arm (c).

parameters. We apply a regular step gradient descent and use the normalized cross correlation (NCC) as similarity measure [3]. Preliminary experiments have shown that all parameters are within a certain range. t_x and t_y are constrained by the image size. s is in the range of $[1.0, 1.4]$, ϕ and θ are in the range of $[-20^\circ, +20^\circ]$ from their given values, α is assumed to have a value of about 0° . Initial values for s , ϕ and θ are taken from image meta information.

The first step initializes the registration in terms of roughly superimposing 2D-MAX and projection image. The second step finds the correct scaling factor and the third step optimizes the projection parameters. The step-wise process was chosen to apply a coarse-to-fine approach.

In the first step, only t_x and t_y are regarded. All other parameters are set to their approximate values. t_x and t_y are initialized with either 1 or 5 different positions: $p_t(1) = \{c\}$, $p_t(5) = \{c, c + p_{x,y}, c + p_{-x,y}, c + p_{x,-y}, c + p_{-x,-y}\}$. c denotes the image center, $c + p_{x,y}$ indicates a point translated towards the corner of the image. MIP₂ is used in this step as it is more robust against initial translations. The maximal NCC value of p_t is regarded as the best translation. The output of the first step is the determination of $t_{x, \text{init}}$ and $t_{y, \text{init}}$.

In the second step, t_x and t_y , s and α are optimized. $t_{x, \text{init}}$ and $t_{y, \text{init}}$ are used for all initializations, α is set to 0. This time, s is varied for different initializations in the assumed range. MIP₁ is used in this step to simulate a 2D-MAX image.

In step 3, ϕ and θ are optimized. Approximate initial values of both parameters are known, t_x and t_y , s and α values are given by the result of the last step. MIP₁ is used as projection technique. In each iteration, ϕ and θ are altered by a value of $\pm\Delta = 5^\circ$ and the other parameters are optimized like in step 2. The optimization process ends if no more improvement in terms of the NCC is measured. All six transformation parameters are determined.

We tested our algorithm on 10 dual-plane 2D-DSA images and 5 corresponding 3D-RA volumes from 5 different patients of the cerebral vasculature. All patients had one or more aneurysms. In one case, the 2D-DSA image was acquired before treatment of the aneurysm, the 3D-RA volume afterwards. The 2D-DSA had a resolution of $[720^2, 2048^2]$ pixels with $[20, 30]$ time frames and a pixel spacing of $[0.08, 0.29]$ mm. The 3D-RA images had a resolution of $[256^2, 512^2]$ pixels with $[149, 444]$ spatial frames and a pixel spacing of $[0.16, 0.53]$ mm.

In our first experiment, we aim to measure the robustness to find the correct initial translation in step 1. The first step is crucial for the whole registration. Starting with manually determined translation parameters as input, we gradually increase the distance of t_x and t_y and measure the capture range by using p_t with either 1 or 5 initialization points and by using MIP₁ or MIP₂. We consider registration step 1 as successful, if the Euclidean distance between computed result and manual estimation of t_x and t_y is lower than 10 mm. This distance is sufficient as initialization for the next step. The capture range is defined as the greatest distance that corresponds to an average success rate of $\geq 75\%$.

The second experiment evaluates the precision of our method by using reference markers. Markers were set at salient image features (bifurcations,

Table 1. Average success rate of step 1 with regard to the number of initialization points (# i) and the used MIP for different distances from the correct position.

# i	MIP	10 mm	20 mm	30 mm	40 mm	50 mm	60 mm	70 mm	80 mm
1	1	100 %	56 %	46 %	52 %	29 %	25 %	31 %	31 %
1	2	96 %	77 %	69 %	48 %	42 %	38 %	38 %	38 %
5	1	100 %	94 %	88 %	75 %	81 %	81 %	69 %	56 %
5	2	100 %	81 %	88 %	100 %	75 %	75 %	69 %	56 %

Table 2. e_{RMS} of the registered test data sets. P₁(2) refers to Patient 1 with angle 2.

measure	P ₁ (1)	P ₁ (2)	P ₂ (1)	P ₂ (2)	P ₃ (1)	P ₃ (2)	P ₄ (1)	P ₄ (2)	P ₅ (1)	P ₅ (2)
e_{RMS} [mm]	1.46	2.22	1.65	2.37	1.58	2.16	1.57	2.35	1.18	1.28

aneurysms) in the acquired image modalities by an expert. We employed a root mean square (RMS) projection error to sum the measured distances between corresponding markers: $e_{\text{RMS}} = \sqrt{\frac{\sum_{i=1}^N (e_i)^2}{N}}$. e_i is the Euclidean distance between corresponding markers and $N = 20$ denotes the number of markers.

3 Results

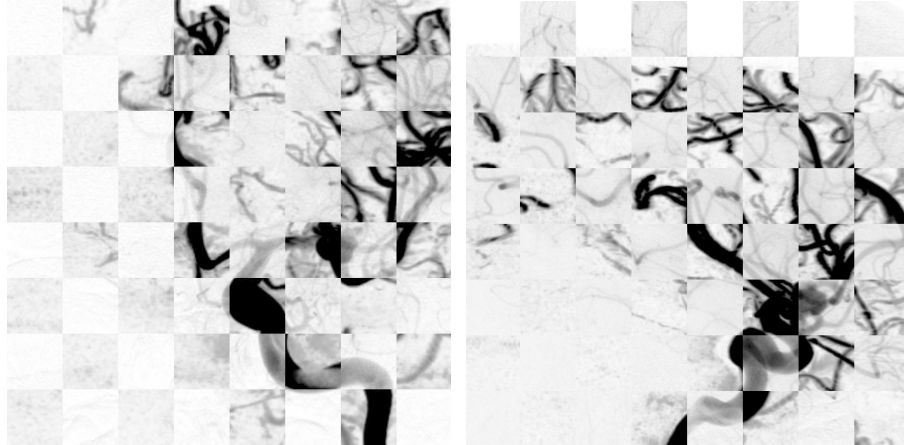
Table 1 summarizes the result of experiment 1. The robustness of step 1 is considerably higher using 5 initialization points than using 1 initialization point. Using a different MIP leads to a different robustness only in case of 1 initialization point. The actual distance in step 1 between starting point and correct determined $t_{x, \text{init}}$ and $t_{y, \text{init}}$ was always smaller than 70 mm in all our test data. The capture range using 5 initialization points is 60 mm in case of MIP₁ and MIP₂. The correct translation parameters could be determined in every case of our test data with 5 initialization points and the standard initialization.

Table 2 shows the result of experiment 2. e_{RMS} is between 1.0 mm and 2.5 mm with an average of 1.78 mm (standard deviation 0.45 mm). Fig. 2 visualizes two example results. The results for images with angle 2 are worse than for angle 1 because of the larger deviation from the initial parameters. We also tested the algorithm with a data set that violates one of our requirements, as an aneurysm treatment was done during image acquisition of 2D-DSA and 3D-RA. The algorithm also produced good results in this case.

4 Discussion

The proposed algorithm computes an automatic 2D/3D-registration. It was tested on a small number of samples. We measured a high robustness and a precision that is in the range of comparable approaches [3] and therefore achieved

Fig. 2. Two examples of registrations of $P_2(1)$ and $P_2(2)$, left and right, respectively. The projection and 2D-MAX image are arranged in a checkerboard pattern. Both images have a different intensity range. Additionally, due to the projection computation the vessel thickness is not always the same in both images. This leads to a challenging visual interpretation of the result images and also hampers the registration process.



our goal. By our three-step strategy we could eliminate the need for user interaction. The registration can be used for validation of simulation results.

Future work will include the extension of the registration algorithm for computing blood flow velocities in DSA data by backprojection.

Acknowledgement. We would like to thank O. Beuing (University Hospital Magdeburg) and S. Serowy (University of Magdeburg) for providing the image data. This work is funded by the LSA (FKZ: 5161AD/0308M).

References

1. Cebal JR, Hernández M, Frangi AF. Computational analysis of blood flow dynamics in cerebral aneurysms from CTA and 3D rotational angiography image data. *Proc Comput Bioeng.* 2003; p. 191–8.
2. Zöllei L, Grimson E, Norbash A. 2D-3D rigid registration of X-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators. In: *Proc IEEE Comput Vis Pat Rec.* vol. 2; 2001. p. 703–969.
3. Hipwell J, Penney G, McLaughlin R, et al. Intensity-based 2-D - 3-D registration of cerebral angiograms. *IEEE Trans Med Imaging.* 2003;22(11):1417–26.
4. Byrne JV, Colominas C, Hipwell J, et al. Assessment of a technique for 2D-3D registration of cerebral intra-arterial angiography. *Br J Radiol.* 2004;77:123–8.
5. Groher M, Zikic D, Navab N. Deformable 2D-3D registration of vascular structures in a one view scenario. *IEEE Trans Med Imaging.* 2009;28(6):847–60.

Adaptive Fokus-Kontext-Kategorisierung für Visualisierungen zur Operationsplanung

Kerstin Kellermann, Alexandra Baer, Bernhard Preim

Institut für Simulation und Graphik, Otto-von-Guericke-Universität Magdeburg
kerstin@isg.cs.uni-magdeburg.de

Kurzfassung. In diesem Beitrag wird ein semantisches Verfahren vorgestellt, mit dessen Hilfe adaptive Visualisierungen zur Therapieplanung generiert werden können. Die dazu notwendige Reduktion wird durch eine Fokus-Kontext-Kategorisierung der im Datensatz enthaltenen wesentlichen Strukturen realisiert. Die Kategorisierung erfolgt anhand einer gewichteten Parameterkombination, die die aktuelle therapeutische Fragestellung repräsentiert. Mit dieser Kategorisierung wird eine adäquate Anpassung der Visualisierung an die spezifische therapeutische Fragestellung ermöglicht. Dieses System ist bei geeigneter Konfiguration in verschiedenen Therapieplanungssystemen einsetzbar.

1 Einleitung

Mit Hilfe interaktiver 3D-Darstellungen der patientenspezifischen Anatomie können verschiedene Fragestellungen geklärt werden, die maßgeblich für die Therapieentscheidung sind. Dazu gehört z.B. der Abstand einer pathologischen Struktur zu einer Risikostruktur. Der Nachteil von 3D-Darstellungen sind die häufig auftretenden gegenseitigen Verdeckungen der Strukturen. Daher ist die Beurteilung der für eine spezifische therapeutische Fragestellung entscheidenden Strukturen und ihrer Relationen zueinander sehr komplex. Eine auf die bzgl. der Fragestellung wesentlichen Strukturen reduzierte Darstellung ermöglicht dagegen eine schnellere Aufmerksamkeitslenkung sowie eine bessere Sichtbarkeit und expressivere Darstellung dieser Strukturen. Eine automatische Anpassung der Strukturparameter beim Wechsel der therapeutischen Fragestellung kann die Exploration der 3D-Darstellung beschleunigen. Zeitaufwändige Vorverarbeitungsschritte einer manuellen Generierung verschiedener aufgabenspezifischer 3D-Darstellungen werden verkürzt bzw. vermieden. Voraussetzung dafür ist die adaptive Bestimmung der wesentlichen Strukturen bzgl. der aktuellen Fragestellung.

Einige Verfahren erzielen eine Hervorhebung der wesentlichen Strukturen mit Hilfe eines Wichtigkeits- bzw. Prioritätswertes pro Struktur [1, 2, 3]. Die damit realisierte Kategorisierung in Fokusobjekte (höchste Wichtigkeit), fokusnahe Objekte (mittlere Wichtigkeit) und Kontextobjekte (geringste Wichtigkeit) [4] optimiert die Sichtbarkeit der wesentlichen Strukturen mittels Smart Visibility Techniken [5]. Die jeweilige Wichtigkeit wird in [1, 3] in einem manuellen Vorverarbeitungsschritt statisch vordefiniert. Dies ermöglicht lediglich einen ansichtabhängigen Blick ins "Innere", der in [3] mit Hilfe eines einfachen Fokusobjektes

(Ebene, Mauszeiger etc.) steuerbar ist. In [2] wird die jeweilige Wichtigkeit entsprechend des selektierten Fokusobjektes angepasst, das die höchste Wichtigkeit erhält, während alle restlichen Strukturen als Kontextstrukturen in den Hintergrund treten. Die Zusammenhänge zwischen den Strukturen hinsichtlich der unterschiedlichen therapeutischen Fragestellungen werden in den genannten Verfahren nicht berücksichtigt. Das Anatomie-Lehrsystem der VoxelMan-Gruppe [6] erfasst systemische, regionale und z.T. auch funktionelle Zusammenhänge in einem umfangreichen semantischen Netzwerk (Ontologie). Darauf basierend werden aufgabenspezifische Visualisierungen im Rahmen der Anatomielehre generiert. Im vorliegenden Beitrag wird ein Verfahren für Interventions- und Operationsplanungen vorgestellt, welches die wesentlichen Strukturen bzgl. der aktuellen therapeutischen Fragestellung automatisch bestimmt. Dabei werden die fallspezifischen geometrischen und pathologischen Relationen in einem Vorverarbeitungsschritt erfasst und in einem parametrisierten Kategorisierungsprozess in Echtzeit ausgewertet.

2 Material und Methoden

In Zusammenarbeit mit HNO-Ärzten der Universitätsklinik in Leipzig wurden zunächst die möglichen therapeutischen Fragestellungen zur Tumor-Operationsplanung inklusive der relevanten Strukturen erfasst. Die Strukturen sind in drei Klassen unterteilbar: pathologische (z.B. Tumore), verdächtige (z.B. Lymphknoten) und anatomische Strukturen (Risikostrukturen, wie Gefäße). Verdächtig bezieht sich hier auf den Verdacht eines pathologischen Charakters bei einer anatomischen Struktur, der noch nicht bestätigt wurde. Diese Strukturen werden bei characterspezifischen Betrachtungen sowohl als pathologisch als auch anatomisch behandelt. Die Lokalisation der pathologischen und verdächtigen Strukturen sowie ihre kritischen minimalen Entfernungen und Infiltrations- bzw. Umschließungsgrade bzgl. der anatomischen Strukturen liegen im Fokus der Operationsplanung. Aus der Analyse der Fragestellungen ergaben sich die notwendigen Parameter, mit denen die wesentlichen Strukturen adaptiv bestimmt werden können [7]. Diese Parameter werden in einem Vorverarbeitungsschritt berechnet und in einer Datenbasis hinterlegt. Neben geometrischen Berechnungen sind hier auch Metainformationen notwendig. Diese beschreiben die Strukturen näher, indem sie z.B. ihre jeweilige Strukturart und Strukturgruppe, aber auch das Stadium eines Tumors oder das Level eines Lymphknotens angeben. Da eine allgemeingültige anatomische Ontologie für sämtliche patientenindividuellen Datensätze zu komplex ist, müssen diese Informationen im jeweiligen Datensatz den entsprechenden Strukturen zugeordnet sein. Zur Unterscheidung der Strukturklassen und zur Bestimmung der kritischsten pathologischen Strukturen werden die Metainformationen Charakter und Schweregrad über interaktiv spezifizierte Klassifizierungsregeln mit folgendem Muster bestimmt:

Charakter : Strukturart : Schweregrad : Messgröße : Vergleich : Grenzwert
Bsp.: pathologisch : Tumor : 3 : max. Durchmesser : > : 40(mm)

Beim Laden des Datensatzes werden dazu für alle Strukturen der hier aufgeführten Strukturarten die notwendigen Messgrößen berechnet. Diese sowie die anwendungsspezifischen Grenzwerte können der jeweiligen TNM-Klassifikation [8] o.ä. entnommen werden. Die den Regeln nicht entsprechenden Strukturen sind anatomisch und haben keinen Schweregrad. Die minimale Entfernung und das Schnittvolumen wird für sämtliche Strukturen anhand ihrer Dreiecksnetze berechnet (z.B. [9]). Zudem werden die Zusammengehörigkeiten der Strukturen bzgl. ihrer Metainformationen erfasst. Der erhöhte Aufwand zur Generierung der Datenbasis hängt hauptsächlich von der Komplexität der geometrischen Berechnungen und der Anzahl der daran beteiligten Strukturen ab. Daher werden diese Ergebnisse im Datensatz dauerhaft abgespeichert.

Der auf diesen akquirierten Parametern basierende Kategorisierungsprozess (Abb. 1) unterteilt die Strukturen in Fokusobjekte, fokusrelevante Objekte und Kontextobjekte und lehnt sich damit der Fokus-Kontext-Definition in [4] an. Die fokusrelevanten Objekte stehen hier jedoch in einem bzgl. der aktuellen therapeutischen Fragestellung wesentlichen Zusammenhang mit den Fokusobjekten. Die Fragestellung wird als eine gewichtete Kombination der im Vorverarbeitungsschritt ermittelten Parameter und einer Menge von selektierten Fokusobjekten ($0-n$) repräsentiert. Zudem werden Bedingungen zur Erfüllung der Parameter mittels bestimmter Werte oder Grenzwerte definiert; z.B. eine minimale Entfernung zu einem Fokusobjekt zwischen 2 und 15 Millimeter oder die Metainformation "Charakter" mit dem Wert "anatomisch". Die Kategorisierung erfolgt schwellenwertbasiert anhand eines Wichtigkeitswertes der jeweiligen Struktur, ähnlich wie in [1]. Dieser berechnet sich durch Aufsummieren der Wichtigkeiten der Parameter ($0-1$), deren Bedingungen diese Struktur erfüllt. Bei den fokusbezogenen Parametern muss dazu lediglich die jeweiligen Bedingung für eines der Fokusobjekte erfüllt sein. Zunächst werden die Strukturen, die den Parametern bzgl. der Metainformationen (z.B. Strukturgruppe = Gefäße) entsprechen, zu Fokusobjekten, wenn ihre derzeitige Wichtigkeit S_{W_1} über dem Schwellenwert t_F liegt (linke Seite von Abb. 1). Erst nach dieser automatischen Erfassung aller Fokusobjekte wird mit den restlichen, den fokusbezogenen Parametern fortgefahren (Abb. 1,

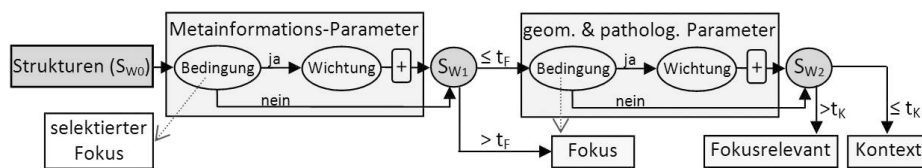


Abb. 1. Die Wichtigkeit einer nicht selektierten Struktur $S_{W_0} = 0$ wird für jeden Parameter, dessen Bedingung sie erfüllt, um die Wichtung des Parameters erhöht. Die Bedingungen beziehen sich dabei z.T. auf die vorhandenen Fokusstrukturen. Strukturen mit einer Wichtigkeit $S_{W_1} > t_F$ werden Fokusstrukturen während die restlichen die geometrischen Parameter und den pathologischen Parameter durchlaufen. Ihre resultierende Wichtigkeit S_{W_2} bestimmt den Schwellenwert t_F entsprechend ob sie fokusrelevant sind oder zum Kontext gehören.

rechter Kasten). Die resultierende Wichtigkeit einer Nicht-Fokusstruktur S_{W_2} bestimmt per Schwellenwert t_K , ob diese fokusrelevant ist oder zum Kontext gehört. Auf diese Weise werden z.B. alle Gefäße zu Fokusstrukturen und alle Primärtumore und Metastasen innerhalb einer minimalen Entfernung von 0 mm bis 5 mm fokusrelevant, während der Rest dem Kontext zugeordnet wird. Zur Einschätzung des Risikos für eine Risikostruktur oder ganzen -gruppe, aber auch dem Risikos durch eine bestimmte pathologische Struktur, werden verschiedene Parameter im pathologischen Parameter kombiniert. Mit diesem werden die Strukturen höher gewichtet, die einen verdächtigen oder anderen Charakter (Metainformation) als die Fokusstruktur haben und innerhalb einer angegebenen kritischen minimalen Entfernung zu dieser Fokusstruktur liegen.

3 Ergebnisse

Es konnte ein Verfahren zur adaptiven Kategorisierung bzgl. wechselnder therapeutischer Fragestellungen für Visualisierungen zur Operationsplanung speziell für den Hals- und Leberbereich entwickelt werden. Dieses schwellenwertbasierte Verfahren arbeitet in Echtzeit und ist robust gegenüber kleinen Änderungen der Parameterwichtungen. Durch Anpassen der Visualisierungsparameter der Strukturen entsprechend ihrer ermittelten Kategorie bzw. Wichtigkeit können die wesentlichen Strukturen hervorgehoben werden. Abbildung 2 zeigt die als wesentlich (Fokus und fokusrelevant) kategorisierten Strukturen eines Halsdatensatzes bzgl. der mit den klinischen Partnern ermittelten Fragestellungen im Vergleich zur Ausgangsdarstellung (a). Abbildung 2(c-e) wurden allein durch den gleich gewichteten pathologischen Parameter für drei verschiedene Fokusobjekte erstellt. Für komplexe Leberdatensätze wurden mit einer annähernd gleichen Parameterkombination vergleichbare Ergebnisse erzielt, die anhand der Bedingungen der Metainformations-Parameter optimierbar sind.

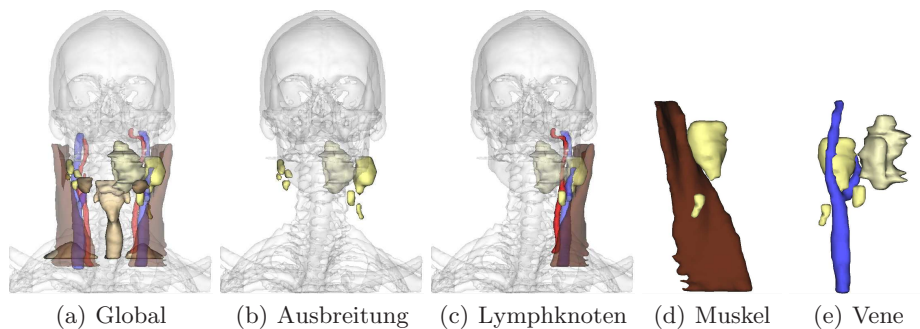


Abb. 2. Als Fokus und Fokusrelevant kategorisierte Strukturen bzgl. der Ausbreitung der pathologischen Strukturen (b), des Risikos durch den vergrößerten Lymphknoten (c) und des Risikos für die Fokusstrukturen in (d) und (e) im Vergleich zur Darstellung sämtlicher Strukturen (a).

4 Diskussion

Mit dem Verfahren wurde eine Grundlage zur adaptiven Generierung aufgabenspezifischer 3D-Visualisierungen geschaffen, auf die z.B. die oben genannten wichtigkeitsgesteuerten Verfahren [1, 3] aufbauen könnten. Das Grundprinzip der Kategorisierung ist in verschiedenen Planungssystemen umsetzbar, wenn die entsprechenden systemspezifischen Parameter verwendet werden. Die in dieser Arbeit verwendeten Parameter sind für die Tumor-Operationsplanung (speziell für Hals und Leber) geeignet. Sie können aber auch für allgemeine Aufgaben herangezogen werden, wie die Selektion der Leber und einen ihrer drei Gefäßbäume. Die exakte Definition der Parameterbedingungen ermöglicht für die Planung eine genaue Extraktion der gesuchten Strukturen. Für die hier entwickelte Anwendung ist eine Dreiecksnetzrepräsentation der Strukturen zur Berechnung geometrischer Parameter und eine geeignete Datenstruktur Voraussetzung. Für die Kategorisierung selbst ist die Dreiecksnetzrepräsentation irrelevant. Das Verfahren wird derzeit um eine einfache Anpassung der Visualisierung erweitert. Dabei soll die Sichtbarkeit der wesentlichen Strukturen und die visuelle Kennzeichnung therapieentscheidender Parameter berücksichtigt werden. Dazu ist auch die Erfassung weiterer Daten vorgesehen, die auf die Kategorisierung selbst keinen Einfluss nehmen. Zum Einblenden einer Auswahl an Orientierungsstrukturen sollen die Kontextstrukturen nach ihrer kontextuellen Relevanz unterteilbar werden. Zudem ist eine Evaluierung geeigneter Visualisierungstechniken geplant, die die Basis für die zukünftige automatische Anpassung der einzelnen Strukturdarstellungen bilden soll.

Literaturverzeichnis

1. Viola I, Kanitsar A, Gröller ME. Importance-driven feature enhancement in volume visualization. *IEEE Trans Vis Comput Graph.* 2005;11(4):408–18.
2. Viola I, Feixas M, Sbert M, et al. Importance-driven focus of attention. *IEEE Trans Vis Comput Graph.* 2006;12(5):933–40.
3. Rautek P, Bruckner S, Gröller ME. Interaction-dependent semantics for illustrative volume rendering. *Comput Graph Forum.* 2008;27(3):847–54.
4. Salah Z, Cunningham D, Straßer W, et al. Perceptually emphasized illustrative visualization for multiple objects. In: *WSI-TechReports WSI-2008-07.* Universitätsbibliothek Tübingen; 2008. p. 1–12.
5. Viola I, Gröller E. Smart visibility in visualization. In: *Comput Aesthet;* 2005. p. 209–16.
6. Höhne KH, Petersik A, Pflesser B, et al. *Voxelman 3D-Navigator: Brain and Skull. Regional, Functional and Radiological Anatomie.* Springer Electronic Media; 2001.
7. Kellermann K. *Ableitung und Verarbeitung semantischer Informationen zur Generierung adaptiver Interventionsplanungs-Visualisierungen [Diplomarbeit].* Otto-von-Guericke-Universität, Fakultät für Informatik. Magdeburg; 2009.
8. Wittekind C, Sobin LH, Klimpfinger M. *TNM-Atlas.* Springer Berlin; 2005.
9. Rössling I, Cyrus C, Dornheim L, et al. Effiziente automatische Bestimmung interventionsrelevanter Entfernungsmaße. In: *Proc BVM;* 2009. p. 66–70.

Detection of Motion Distorted Areas in Perfusion MRI of the Breast

Sebastian Schäfer, Klaus D. Tönnies

Department of Simulation and Graphics, University of Magdeburg
schaefer@isg.cs.uni-magdeburg.de

Abstract. Dynamic contrast-enhanced MRI of the breast (DCE-MRI) is widely used for detection and quantification purposes of breast cancer. The particular acquisition procedure of DCE-MRI over time leads to motion artifacts which distort analysis results. In this work we examine two criteria for a segmentation approach to separate areas affected by motion artifacts from properly perfused regions. By application of our approach to simulated perfusion data we showed that a homogeneity measure based on perfusion parameters achieves the best segmentation results with respect to the ground truth. This effect was confirmed when segmenting patient data.

1 Introduction

In North America and Europe, breast cancer is the most frequent mortal disease in women. Medical imaging techniques play an important role in diagnosis. In particular high-risk cases DCE-MRI of the breast allows a sophisticated analysis of lesion dynamics. A breast tumor leads to formation of new vessels, which is referred to as angiogenesis [1]. In DCE-MRI a contrast agent (CA) is intravenously injected and works as a tracer of perfusion. The angiogenetic activity leads to CA accumulations and allows for breast cancer detection through the calculation of perfusion parameters.

A current method to determine perfusion parameters for different lesions is to define regions with similar perfusion characteristics (Regions of Interest, ROI). Calculating an average over the ROI reduces the influence of noise. Many tools allow for manual ROI placement, although this carries the risk of missing important details of the data and is subject to inter- and intra-observer variability. We employ a method by Glaßer et al. [2] to automatically generate ROIs in DCE-MRI to derive perfusion parameters of tumorous tissue.

Motion artifacts which are based on patient motion during acquisition hamper the correct determination of perfusion parameters. They result in incorrect inter-voxel correspondences between images at different times. In the automatic region generation method, this results in smaller regions of homogeneously perfused tissue and produces new regions with mixed perfusion parameters.

To reduce the influence of motion a global registration approach can be applied (e.g. we use a method by Rueckert et al. [3]). It optimizes the voxel po-

sitions using mutual information as similarity measure and a continuous deformation function (e.g. B-Splines). Hence, there exists a high conformance at the boundary between breast and background but in tumor regions within the breast it is less accurate. This is caused by the fact that entropy is high even for images in perfect registration because of change of CA concentration. This leads to a less well pronounced minimum for the registration criterion. Other approaches have tried to compensate for this. Melbourne et al. [4] analyze principal components of the currently best registered two images and optimize them in an iterative process. If periodic motion is present though, it will be propagated in the principal components. Xiaohua et al. [5] merge the process of segmentation and registration to incorporate knowledge of regional anatomy to align different images, but no reliable evaluation for non-rigid registration results is provided.

Although many attempts have been conducted, a perfect registration is still hard to achieve. Thus, remaining motion artifacts have to be identified by abnormal perfusion parameter characteristics. The goal of this study is to examine which criteria are well suited for this purpose.

2 Materials and Methods

For the investigation presented here we use a region merging segmentation approach [2] and tested two different merging criteria. These criteria are defined such that the algorithm produces homogeneous regions in terms of a defined correlation threshold. The first criterion employs the raw intensity values which have been measured at particular time steps. As regions contain multiple voxels, image values are averaged over each time measurement when merged with other regions. The vector V_1 containing relevant information for a region R is

$$V_1(R) = (R_0, \dots, R_{n-1})^T \quad (1)$$

where n is the number of time points measured in the dataset. It is assumed that similar perfusion characteristics lead to similar measured values in two different regions. It should be noted that image artifacts, particularly noise, motion and partial volume effects will be directly represented in the vector data. Hence, they will influence the calculation of similarity of regions.

The second criterion uses general properties of the averaged perfusion time course of a region. The curve properties are employed as indicators for diagnostic purposes as well [6]. The feature vector V_2 in ((2)) incorporates three parameters which are calculated from the perfusion time curve.

$$V_2(R) = (S_{up}(R), S_{down}(R), RE_m(R))^T \quad (2)$$

$$S_{up}(R) = \left(\frac{R_{\lfloor n/2 \rfloor} - R_0}{R_0} \right) \quad (3)$$

$$S_{down}(R) = \left(\frac{R_{\lfloor n/2 \rfloor} - R_{n-1}}{R_{n-1}} \right) \quad (4)$$

$$RE_m(R) = \left(\frac{\max(R_0, \dots, R_{n-1}) - R_0}{R_0} \right) \quad (5)$$

Table 1. Comparison of criterion V_1 and V_2 using different statistical measures. The best performing measures (SP, SE, REL, NPV) of either V_1 or V_2 are printed bold.

displacement	criterion V_1				criterion V_2			
	SP	SE	REL	NPV	SP	SE	REL	NPV
1 pixel	0.912	0.878	0.979	0.612	0.876	0.905	0.970	0.675
2 pixel	0.983	0.851	0.980	0.872	0.986	0.934	0.986	0.932
3 pixel	0.972	0.827	0.965	0.855	0.974	0.876	0.969	0.895

The upslope (Eq. (3)) characterizes the early enhancement intensity, the downslope (Eq. (4)) shows the washout of CA from the tissue and the relative enhancement in ((5)) represents the maximum enhancement of the curve. R denotes the current region and n the number of measured time steps.

Two different experiments have been accomplished to test both similarity criteria for their ability to distinguish between normal and motion distorted areas in the image. The first experiment is performed with simulated data, showing a lesion with two differently enhancing regions. Artificial motion displacement is added with different magnitude and relative enhancement curves (REC) of segmented regions are evaluated. This shows the effects solely caused by motion artifacts. The second experiment is performed with patient data (manually registered and unregistered) to see whether similar effects as in the first experiment can be observed.

3 Results

In this section we describe the outcome of the experiments previously outlined. The different rows in Fig. 1 show the results of the segmentation of the simulated dataset. The extent of motion influence increases from 0 pixels to up to 3 pixels offset per time step. The 1 pixel and 2 pixel offset is a linear shift whereas the 3 pixel image is generated by a rotation.

Tab. 1 shows the detection performance for both criteria of the simulated images. Therefore all segments have been manually classified as regions representing motion artifacts (M) and regions not influenced by motion (P). The specificity (SP) expresses the likelihood of a segment being classified as P and belonging to class P. The sensitivity (SE) expresses the likelihood of a segment being classified as M and belonging to class M. The relevance (REL) expresses the proportion of segments correctly classified as P and all segments classified as P. The negative predictive value (NPV) expresses the proportion of segments correctly classified as M and all segments classified as M.

For the second experiment a ROI of a DCE-MRI data set showing contrast enhancement in tumorous tissue has been manually registered by an expert. Both the registered and unregistered image have been segmented with the proposed approach. Results are shown for the parameter-based criterion V_2 (Fig. 2), because it achieves better results in the first experiment.

4 Discussion

Both similarity criteria are able to segment the two differently perfused areas of the simulated data (first row of Fig. 1). However, using criterion V_2 results in much fewer regions (avg. of 20) compared to criterion V_1 (avg. of 40). This fact enables a better distinction between regions showing correct perfusion characteristics and those influenced by motion (see column 3 and 5 in Fig. 1). In Tab. 1 SE and the NPV characterize the performance of segmenting motion distorted areas. The criterion using V_2 leads to a better performance, especially when a stronger influence of motion is present. V_2 also performs better for the segmentation of correctly perfused areas (SP and REL). Therefore this criterion should be preferred. The second experiment analyzes patient data. The regional REC of the manually registered image are very similar and indicate normal perfusion

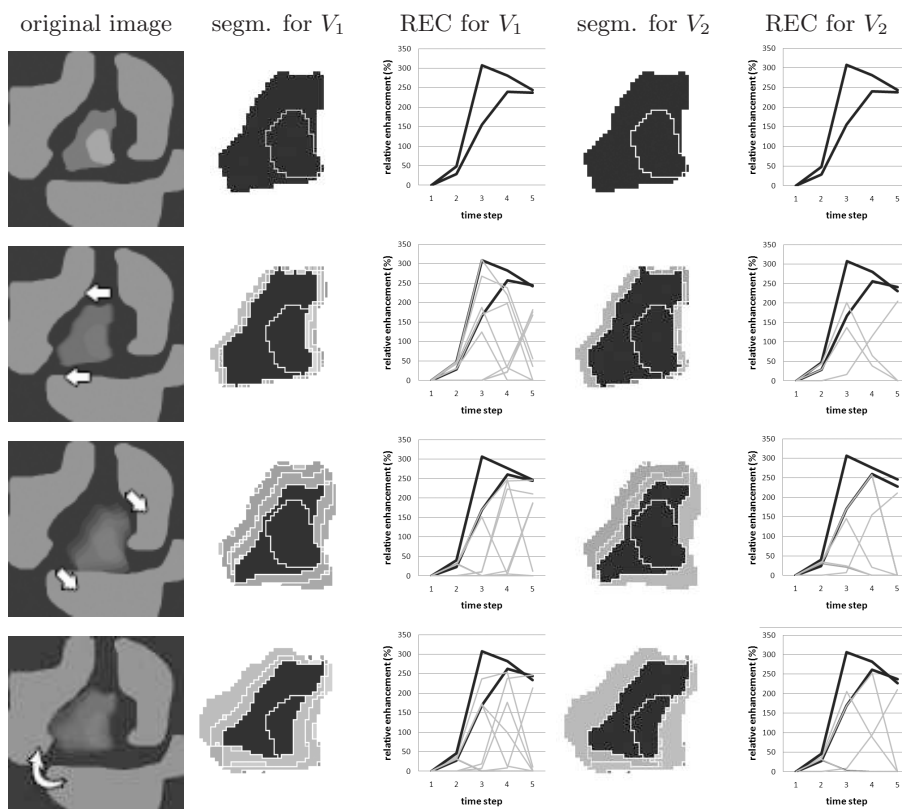
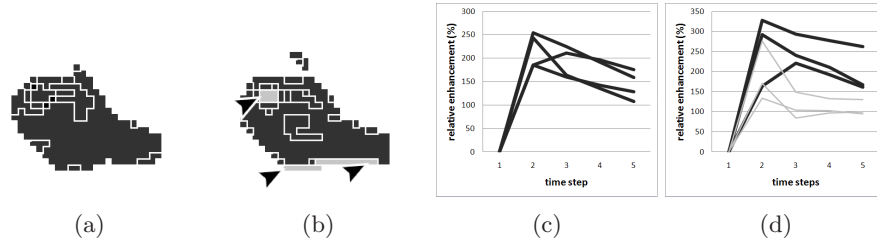


Fig. 1. Each row depicts the results of simulated perfusion data with different extent of motion influence. The first column shows the image itself, the second shows the segmentation result for criterion V_1 and the third column the enhancement curves for the segmented regions. Column 4 and 5 show the results for criterion V_2 likewise. The two region curves showing the true perfusion characteristics are printed bold and black.

Fig. 2. Segmentation results for manually registered (a) and unregistered (b) patient data is shown. (c) and (d) present the REC of (a)/(b) for regions bigger than 5 voxels.



characteristics (Fig. 2). The unregistered image leads to formation of new regions near the boundary (black mark) which show different time courses (dotted curves) than the correctly perfused parts. This implies that similar effects than in the first experiment can be observed in patient data.

Further steps will be focused on the motion correction to DCE-MRI. The proposed method can then be used to evaluate and conduct the correction process although more experiments have to be performed with patient data. As the parameter-based model yields promising results, the use of a more sophisticated model might be useful. Though, we plan to implement a pharmacokinetic model to measure the plausibility of derived REC to automatically distinguish between realistic perfusion parameters and those influenced by artifacts.

Acknowledgement. We thank Uta Preim for providing the medical data. This work has been funded by DFG (research grant no. TO166/13-1).

References

1. Kuhl CK. The current status of breast MR imaging, Part I. *Radiology*. 2007;244(2):356–78.
2. Glaßer S, Schäfer S, Oeltze S, et al. A visual analytics approach to diagnosis of breast DCE-MRI data. *Proc VMV*. 2009; p. 351–62.
3. Rueckert D, Sonoda L, Hayes C, et al. Nonrigid registration using free-form deformations. *IEEE Trans Med Imaging*. 1999;18(8):712–21.
4. Melbourne A, Atkinson D, White M, et al. Registration of dynamic contrast-enhanced MRI using a progressive principal component registration (PPCR). *Phys Med Biol*. 2007;52(17):5147–56.
5. Xiaohua C, Brady M, Rueckert D. Simultaneous segmentation and registration for medical image. In: *Proc MICCAI*; 2004. p. 663–70.
6. Preim B, Oeltze S, Mlejnek M, et al. Survey of the visual exploration and analysis of perfusion data. *IEEE Trans Vis Comput Graph*. 2009;15(2):205–20.

Improvement and Evaluation of a Time-of-Flight-based Patient Positioning System

Simon Placht, Christian Schaller, Michael Balda, André Adelt,
Christian Ulrich, Joachim Hornegger

Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg
`christian.schaller@informatik.uni-erlangen.de`

Abstract. In this paper we improve a surface-based patient positioning method. The method describes a system for automatic positioning of patients using time-of-flight cameras in radiotherapy. To improve the registration result three new preprocessing steps (bilateral filtering, temporal averaging, variance filtering) are introduced to the processing pipeline. Furthermore, the accuracy of the surface matching algorithm (ICP) is improved by changing the distance measurement from point-to-point to point-to-surface. The mean registration error is improved by 2.14 mm to 0.74 mm, whereby the working distance could also be increased from 0.8 m to 1.5 m.

1 Introduction

The precise delivery of irradiation to the target volume within radiotherapy and particle therapy depends heavily on the correct positioning of the patient. In the worst case dose delivery to a false location can harm healthy tissue [1]. Therefore, accurate patient positioning in radiotherapy and particle therapy is an important issue for high precision treatment.

Both the treatment plan coordinate system and the patient coordinate system have to be aligned properly before each treatment session. A pre-treatment CT is acquired and used as a basis for the treatment plan. Within this CT dataset the shape and position of the tumor can be verified and a dose delivery plan can be created. Right before the treatment session starts, the tumor position has to be aligned with the isocenter of the linear accelerator. Automatic patient positioning also increases the patient throughput [2]. Surface-based patient positioning is a convenient way to position patients without radiation.

A commercial available surface based system for patient positioning is the VisionRT system (www.visionrt.com). Schoeffel et al. [3] investigated this system and obtained an accuracy of $0.40 \text{ mm} \pm 0.26 \text{ mm}$ for rigid phantoms. In this paper we use a similar evaluation setup.

Schaller et al. [4] proposed an approach where a time-of-flight (TOF) camera is used to acquire the 3-D shape of a patient in order to position it with respect to a priorly acquired reference surface. Compared to the VisionRT system,

such a system uses only one camera and is much more cost-effective. They could achieve a mean registration error of 2.88 mm for rigid body phantoms at a camera distance of 80 cm.

2 Materials and Methods

For testing and evaluation we used a rigid plaster cast body phantom, which is described in detail in [5]. Utilizing this phantom, a simulation close to clinical conditions can be achieved. Furthermore, a recent TOF camera, the CamCube from PMD Technologies is used. Details about TOF cameras can be found in [6].

Due to the measuring principle of TOF cameras the data provided by this modality is suffering from various problems (Fig. 1). The most prominent problems regarding surface registration are the high noise level of the data and so called flying pixels. These occur whenever a distance discontinuity between two objects in the field of view of the TOF camera is given. Both problems make it difficult to register two surfaces acquired by a TOF camera in a robust and absolute manner. Due to the ToF sensor matrix, which is 204×204 pixels, the computed 3D coordinates exactly match a grid structure in x and y direction. A standard iterative closest points (ICP) algorithm [7] using an Euclidean point-to-point distance measurement prones to suffer from a “snap-to-grid” effect. For

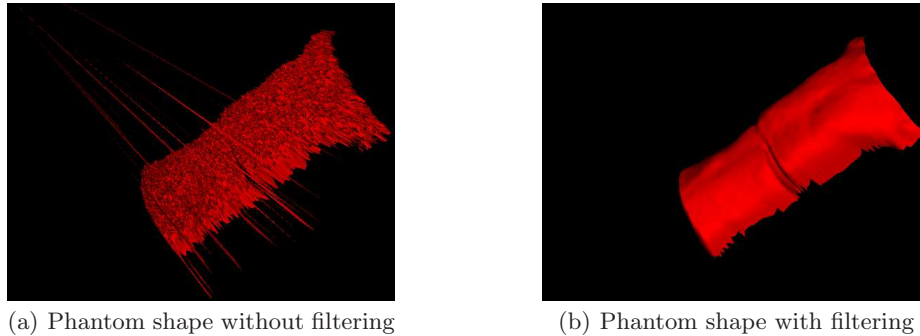


Fig. 1. Comparison of preprocessing steps.

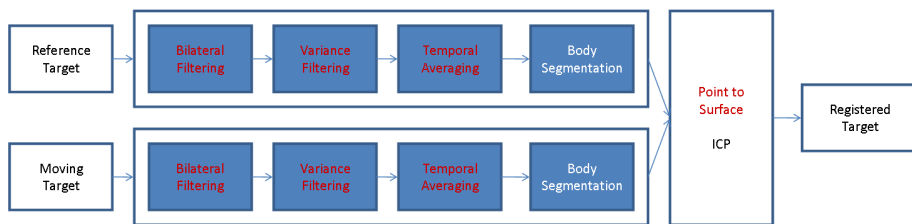


Fig. 2. Processing pipeline, showing the standard components used by Schaller et al. [4] and improved/new components (in red) to obtain much more stable results.

a distance of 1.0 m the grid spacing is already about 3.5 mm. To improve the registration result all crucial components of the registration pipeline are modified (Fig. 2).

First of all, the noisy raw data need to be preprocessed. A set of filters is applied to reduce the temporal as well as the spatial noise artifacts. In order to cope with spatial noise, bilateral filtering [8, 9] is applied to the distance matrix. Since bilateral filtering is an edge preserving smoothing technique, object structures can be preserved much better than with the Gaussian filter utilized in [4]. Temporal noise is suppressed by averaging the 3D information of the camera over a certain amount of frames. Furthermore, we introduce a variance filter to eliminate bad pixels. For each pixel, we compute the variance based on the last frames. By applying a threshold to these values, “flying pixels” can be eliminated.

As a further modification to improve the registration accuracy, we changed the strategy to find point correspondencies in the ICP algorithm from point-to-point to point-to-surface.

Figure 3 shows the test setup, which is similar to [3]. The movement of a Siemens ONCOR patient table was observed with three gauges – one for each axis. The gauges used for this evaluation had a total measurement range of 1 cm and an accuracy of $10\ \mu\text{m}$. The TOF camera was perpendicularly aligned to the table in a distance of about 1.5 m, which is a suitable distance for a practical setup. Instead of a treatment couch the authors of [4] use a robot arm with a precision of 0.1 mm to position a phantom. For the evaluation the patient table was moved arbitrarily in one direction with the other two directions locked. The evaluated relative table translations were in a range between 0.5 mm and 9.5 mm, which is very similar to the experiments described in [3]. In contrast to [3] and [4], rotations as well as translations in three dimensions at the same time were not considered within these experiments. The translation provided as output of the ICP algorithm was compared to the real translation given by the gauges.

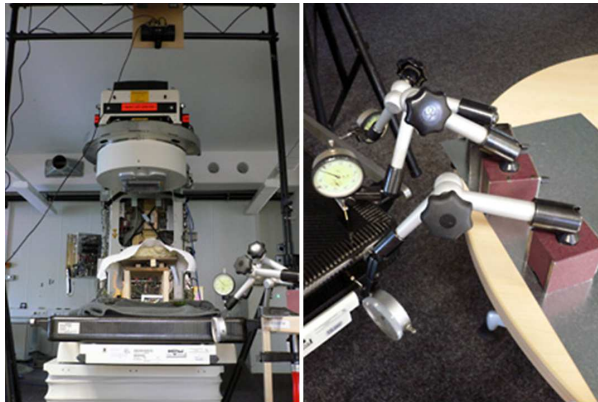


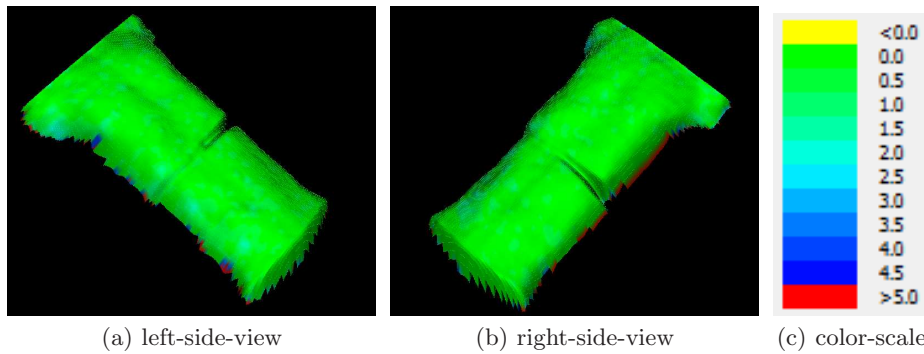
Fig. 3. Test setup: left side: Siemens ONCOR system with ToF camera viewing the body plaster cast phantom; right side: three measurement gauges to measure patient table displacement with an accuracy of 1/100 mm.

Table 1. RMS of the Euclidean error of the computed translation with respect to the ground truth for lateral (x), longitudinal (y) and vertical (z) displacements.

ground truth	min [mm]	max [mm]	mean [mm]	median [mm]	std [mm]
0.50 mm translation in x	0.17	0.39	0.30	0.30	0.17
0.95 mm translation in x	0.28	0.56	0.42	0.40	0.24
2.50 mm translation in x	0.26	0.41	0.35	0.36	0.14
5.35 mm translation in x	0.20	0.63	0.47	0.49	0.32
6.45 mm translation in x	0.37	1.20	0.77	0.62	0.66
7.15 mm translation in x	0.30	0.67	0.51	0.51	0.30
9.50 mm translation in x	1.64	1.88	1.75	1.75	0.49
0.50 mm translation in y	0.00	0.42	0.22	0.17	0.24
5.15 mm translation in y	1.03	1.43	1.26	1.20	0.56
7.15 mm translation in y	0.30	0.93	0.54	0.42	0.49
9.15 mm translation in y	0.81	1.46	1.20	1.17	0.66
0.95 mm translation in z	0.10	0.55	0.39	0.42	0.32
2.95 mm translation in z	0.45	0.66	0.54	0.53	0.24
6.35 mm translation in z	1.08	1.24	1.16	1.16	0.32
7.60 mm translation in z	1.14	1.37	1.23	1.22	0.37
average	0.54	0.92	0.74	0.71	0.37

3 Results

Table 1 shows the RMS error of the Euclidean distance between ground truth and computed result. The statistical indicators mean, median and standard deviation are thereby determined upon a sequence of 50 consecutive frames of the non-moving phantom. Figure 4 illustrates the surface distance between the reference and the transformed source dataset.

**Fig. 4.** Exemplary registration result: Color coded distance between transformed source and target dataset (range: 0 mm – 5 mm).

In comparison to [4], we could decrease the mean positioning error from $2.88 \text{ mm} \pm 1.84 \text{ mm}$ to $0.74 \text{ mm} \pm 0.37 \text{ mm}$, although the distance between camera and treatment couch was incremented from 0.80 m to 1.50 m. The proposed modified algorithm is capable to process with 1 fps on a standard dual-core 2.0 GHz CPU at the moment.

4 Discussion

In the future, the algorithm has to be improved regarding runtime and robustness for displacements higher than 1 cm. Currently, a feature-based pre-registration algorithm is investigated, which can provide a good initial translation and rotation of the body shape. This can be used as initialization for the ICP algorithm to refine the registration result.

References

1. Moore CJ, Graham PA. 3D dynamic body surface sensing and CT body matching: a tool for patient set-up and monitoring in radiotherapy. *Comput Aided Surg.* 2000;5(4):234–45.
2. Sweeney R, Vogele M, Wegmayr A, et al. The patient positioning concept for the planned MedAustron centre. *Radiother Oncol.* 2004 12;73(2):64–7.
3. Schoffel PJ, Harms W, Sroka-Perez G, et al. Accuracy of a commercial optical 3D surface imaging system for realignment of patients for radiotherapy of the thorax. *Phys Med Biol.* 2007;52:3949–63(15).
4. Schaller C, Adelt A, Penne J, et al. Time-of-flight sensor for patient positioning. In: *Proc SPIE.* vol. 7258; 2009.
5. Ulrich C, Schaller C, Penne J, et al. Evaluation of a time-of-flight based respiratory motion management system. In: *Proc BVM;* 2010. p. to appear.
6. Xu Z, Schwarte R, Heinol H, et al. Smart pixel: photometric mixer device (PMD): new system concept of a 3D-imaging-on-a-chip. In: *Proc Int Conf Mechatronics and Machine Vision in Practice;* 1998. p. 259–64.
7. Rusinkiewicz S, Levoy M. Efficient variants of the ICP algorithm. In: *Proc 3DIM;* 2001. p. 145–52.
8. Tomasi C, Manduchi R. Bilateral filtering for gray and color images. In: *Proc ICCV;* 1998. p. 839.
9. Paris S, Durand F. A fast approximation of the bilateral filter using a signal processing approach. *Int J Comput Vision.* 2009;81(1):24–52.

Einfluss von Formvariationen auf Finite Elemente Simulationen bei muskulären Strukturen

Lars Walczak¹, Frank Weichert¹, Andreas Schröder², Constantin Landes³,
Heinrich Müller¹, Mathias Wagner⁴

¹Lehrstuhl für Graphische Systeme, Technische Universität Dortmund

²Institut für Mathematik, Humboldt Universität zu Berlin

³Kiefer- und Plastische Gesichtschirurgie, Universitätsklinikum Frankfurt

⁴Institut für Allgemeine und Spezielle Pathologie, Universität des Saarlandes

`lars.walczak@tu-dortmund.de`

Kurzfassung. Heutzutage existieren konkurrierende Konzepte zur chirurgischen Behandlung von Fehlbildungen im Lippen-, Kiefer- und Gaumenbereich. Zur Unterstützung von Chirurgen bei der Vorhersage von Operationsergebnissen zählen morphologiebasierte Finite Elemente Simulationen in histologischer Auflösungsgenauigkeit zu den vielversprechendsten Methoden. Dabei stellt die Erstellung von geeigneten Netzen, die sich gleichermaßen zur mathematischen Modellierung sowie zur genauen Rekonstruktion der Anatomie eignen, eine unzufriedenstellend gelöste Aufgabe dar. Hier wird daher untersucht, wie sich Formvariationen auf numerische Ergebnisse auswirken. Die Essenz dabei ist, dass eine exakte Segmentierung eine unabdingbare Voraussetzung ist, aber auch kleinere Abweichungen in der Formdarstellung zu größeren Änderungen im Simulationsergebnis führen können.

1 Einleitung

Aktuelle Entwicklungen bei computergestützten Simulationsmethoden in der medizinischen Praxis - beispielsweise zur Operationsplanung - basieren üblicherweise auf der Kombination detaillierter Anatomieinformation aufgrund segmentierter Bilder, der dreidimensionalen Rekonstruktion des zu untersuchenden Gebiets sowie der numerischen Simulation des Deformationsverhaltens. Diese Studie diskutiert den Einfluss verschiedener Segmentierungen aufgrund unterschiedlicher Beobachter am Beispiel eines quergestreiften Gesichtsmuskels, des ringförmigen Musculus Orbicularis Oris (MOO). Dieser stellt einen Ansatzpunkt bei der Behandlung von Fehlbildungen im Lippen-, Kiefer- und Gaumenbereich dar [1]. Die Mikromorphologie hat einen nicht zu vernachlässigenden Einfluss auf jedes korrektive Konzept, beispielsweise bei der Korrektur der Lippe. Gängige Vorgehensweisen tendieren dazu, ein mögliches mikromorphologisch korrektes Modellieren der muskulären Strukturen zu Gunsten einer detaillierten Hautrekonstruktion zu vernachlässigen. Diese erlaubt zwar Symmetrie im Ruhezustand

der Gesichtsmuskulatur, aber ersichtliche Asymmetrie bei Lippenbewegungen. Computerunterstützte mathematische Modellierung kann einen wichtigen Beitrag zur Analyse der resultierenden Deformationsvektoren bei Muskelaktivität leisten, denn postoperative Resultate unterschiedlicher Rekonstruktionsmethoden können so reproduzierbar verglichen werden. Eine natürlichere postoperative Kraftverteilung in silico könnte einen Chirurgen in die Lage versetzen, harmonischere und natürlich funktionale Lippenbewegungen in vivo zu ermöglichen. Exakte, manuelle Segmentierungen histologischer Großflächenschnittpräparate [2], wie sie in diesem Fall vorliegen, führen zu hoch komplexen Polygonzügen, deren 3D-Rekonstruktion sich schwierig gestaltet. Die Frage am Beispiel des MOO ist, ob der durch manuelle Segmentierung erzeugte Detailgrad für die Berechnungen überhaupt notwendig ist, bzw. welcher Detailgrad angemessen ist. Vor diesem Hintergrund soll herausgefunden werden, bis zu welchem Grad eine Netzrepräsentation [3] des MOO die anatomischen Beobachtungen in einer Finite Elemente (FE) Simulation wiedergeben muss.

2 Material und Methoden

Zur Bearbeitung der einleitend motivierten Problemstellung werden die originalen Polygonzüge zur Variation derart geglättet, dass graduell hohe Frequenzen entfernt werden, bis lediglich eine Grundstruktur des Objekts erhalten bleibt. Das hierzu genutzte Verfahren sind die bekannten Fourier Deskriptoren [4]. Die Ausgangsdaten und deren Modifikationen werden nach 3D-Rekonstruktion in Form von Netzen miteinander verglichen, dabei werden ebenso unterschiedliche Netzauflösungsstufen betrachtet.

Korrekte 3D-Rekonstruktionen benötigen ein Netz, welches die Geometrie und Topologie einer Vorlage entsprechend approximiert. Dieser Vorgang gestaltet sich schwierig, wenn die Konturen komplexe Geometrien aufweisen und mehr als einen Gegenpart auf einer benachbarten Schnittebene besitzen. Der folgende, kurz skizzierte Ansatz (Details z.B. in [2]) löst diese Probleme elegant, bestimmt er doch eine Approximation der manuell segmentierten Konturen über einen Level Set, welcher eine implizite Oberfläche ϕ darstellt. Über einen zeitlichen Evolutionsprozess wird ausgehend von einer initialen Schätzung für ϕ anhand der Level Set Gleichung $\frac{\partial \phi}{\partial t} = \nabla d(\mathbf{x}) \cdot \nabla \phi$ die Oberfläche ϕ bestimmt, wobei d ein aus den Polygonen bestimmtes Distanzfeld darstellt. Um ein passendes, FE-konformes Tetraedernetz zu generieren, wird die implizite Repräsentation ϕ dazu genutzt, um einen iterativen Octree-Unterteilungsprozess zu steuern. Zur Approximationsverbesserung wird das resultierende Tetraedernetz auf das Nullniveau von ϕ deformiert. Eine Netzqualitätsanalyse deutet auf eine gute Eignung für das hier vorliegende Problem hin [2].

Unter gewissen Vereinfachungen stellen hyperelastische Materialgesetze einen physikalisch begründbaren Rahmen und damit ein angemessenes Modell für biologisches Weichgewebe auf einer makroskopischen Skala dar. Die meistgenutzten Ansätze basieren auf transversen, isotropen, hyperelastischen Materialgesetzen, wie etwa dem von Mooney-Rivlin [5]. Um die Gesamtkomplexität gering zu

Tabelle 1. Vergleich der erzeugten FE Netze. Anzahl der Unbekannten und Konditionszahl der Systemmatrix für MOO und Referenztorus. Datensatz kennzeichnet den jeweiligen Untersucher, original bezeichnet unveränderte Segmentierungen, glatt steht für modifizierte Varianten.

Datensatz	Segment	Unbek., fein	Kondition, fein	Unbek., grob	Kondition, grob
“A”	original	117.978	3, 14035 <i>E7</i>	23.418	3, 39126 <i>E6</i>
“A”	glatt	112.656	6, 92955 <i>E7</i>	24.027	5, 41920 <i>E6</i>
“C”	original	77.202	8, 60515 <i>E6</i>	15.699	7, 82299 <i>E5</i>
“C”	glatt	80.064	2, 41594 <i>E6</i>	15.993	1, 24151 <i>E5</i>
“G”	original	103.257	2, 64039 <i>E7</i>	20.373	1, 07328 <i>E6</i>
“G”	glatt	100.608	6, 21062 <i>E7</i>	19.401	9, 23560 <i>E5</i>
Torus	original	51.474	1.138.141	11.898	241.679

halten, werden keine Muskelfaserrichtungen oder Aktivierungen zur Muskelkontraktion berücksichtigt und lediglich das passive Deformationsverhalten, welches durch eine idealisierte Volumenkraft induziert wird, analysiert. Die Verbindung des Muskels zu anderen Strukturen, wie beispielsweise Knorpeln oder Knochen, wird über Dirichlet-Randbedingungen formuliert.

Die hier vorliegende nichtlineare variationelle Formulierung des Modells wird mit Hilfe eines verschiebungsbasierten, stückweise linearen FE Ansatzes gelöst, um Deformationen sowie Spannungen numerisch zu bestimmen. Als Hauptspannung bezeichnet man den größten der drei Eigenwerte des symmetrischen Cauchy'schen Spannungstensors.

3 Ergebnisse

Im Hinblick auf die Validierung der aufgezeigten Methoden wurde der Musculus Orbicularis Oris rekonstruiert, einmal auf Basis von drei originalen (manuellen) Segmentierungen sowie in geglätteten Varianten. Zudem wurden je zwei Netzauflösungsstufen untersucht (grob, fein). Der Einfluss der Formen und damit zusammenhängend der Detailgrad stand im Mittelpunkt. Um die Ergebnisse der hyperelastischen Deformation und der Netzgenerierung besser einordnen zu können, wurde ein Referenzmodell in Form eines Torus als idealisierter MOO herangezogen. Die aus dessen Simulation resultierenden Hauptspannungen sind in Abb. 1(a) dargestellt. Wie erwartet zeigen sich symmetrische Deformationen.

In den Abb. 1(b-d) werden die Simulationsergebnisse in Form von Hauptspannungen für den MOO dargestellt. Die Spannungsmaxima konzentrieren sich analog zum Torus dort, wo der Muskel mit seiner Umgebung verbunden ist. Grobe Netze können die Level Set Oberfläche nicht genau approximieren, da die Tetraeder zu groß sind und feine Details nicht abgebildet werden können.

4 Diskussion

Die Hauptspannungskurven in Abb. 1 zeigen die typischen Verläufe der Tori mit einer Ausnahme nur in den feinaufgelösten Netzen des MOO. Daher muss eine bestimmte Netzaufösung und damit eine detaillierte Segmentierung gegeben sein, um sinnvolle medizinische Erkenntnisse zu gewinnen. Dies spricht sehr für den ausgewählten, hochaufgelösten Ansatz auf Basis der histologischen Schnittpräparate. Die Beträge der Spannungen liegen höher als dies beim Torus der Fall war. Dieser Aspekt steht im Zusammenhang mit der differierenden Geometrie zwischen MOO und Torus. Eine signifikante Auffälligkeit manifestiert sich im untypischen Kurvenverlauf für den Datensatz "C", dessen Ausprägung durch teilweise schwierig zu segmentierende anatomische Strukturen bedingt ist. Daher ergeben sich insbesondere Gemeinsamkeiten oder Unterschiede in der Segmentierung als relevante Aspekte für weitere Untersuchungen. Eine Analyse der Konditionszahlen der Systemmatrizen (Tab. 1) sowie der Elementqualität führt zu keinem abschließenden Ergebnis, da sich die Statistiken aller generierten Netze für den MOO nicht signifikant unterscheiden. Die Geometrie und damit die Segmentierung haben daher einen großen Einfluss auf das Ergebnis – dieses spricht für die Notwendigkeit möglichst sorgfältiger Segmentierungen.

In der vorliegenden Studie wurde zunächst anhand eines Netzgenerierungsalgorithmus, welcher hohe physische und visuelle Exaktheit aufweist, der Einfluss

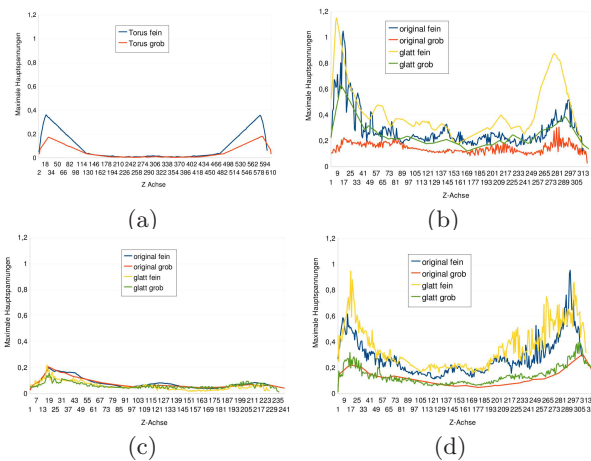


Abb. 1. (a) Simulationsergebnisse für den Torus, fein (blau, 64.679 Tetraeder), grob (rot, 14.703 Tetraeder). Die FE Simulation erzeugt nahezu symmetrische Spannungsverläufe. (b-d) MOO: Hauptspannungen für Netze von verschiedenen Untersuchern. Jedes Segment (original, geglättete Variante) wurde mit zwei unterschiedlichen Netzaufösungen approximiert (grob ca. 20-30 tsd. Tetraedern, fein ca. 100-150 tsd. Tetraedern). Lediglich feine Netze weisen Spannungsspitzen an den festgehaltenen Rändern auf, grob aufgelöste Netze weisen relativ flache Kurven im direkten Vergleich auf. (b) Datensatz "A", (c) Datensatz "C", (d) Datensatz "G".

der Segmentierung oder Glättung derselben auf die numerischen Ergebnisse untersucht. Unter Beachtung der medizinischen Implikationen nach einer zukünftigen Verifikation der Resultate, sollte eine Vorabsimulation den Operateur in die Lage versetzen zu entscheiden, ob ein chirurgischer Eingriff wie vorgesehen möglich ist oder besser vermieden werden sollte, etwa aufgrund farbkodierter Spannungsmuster [2]. Ein bestimmter Detailgrad der Segmentierung sollte dabei eingehalten werden, um Spannungsspitzen korrekt abzubilden. Dies fällt Untersuchern nicht immer leicht. So konnten in einigen Schnittebenen anatomische Strukturen nur schwer voneinander unterschieden werden (Ergebnisse zu Datensatz "C"). Zudem weist der mechanische Erstellungsprozess bei der Materialpräparation inhärente Verzerrungen auf. Diese Artefakte tragen bedingt an Unsicherheit bei der Erstellung eines exakten Computermodells bei. Die hier vorliegenden quantitativen Daten sind daher nur bedingt mit Angaben in der Literatur vergleichbar. Allerdings demonstriert diese Studie, dass eine Tetraedernetzapproximation ausreichen könnte, um eine realistische Muskel-Simulation mittels der FE Technik unter gewissen Voraussetzungen durchzuführen.

Die vorliegenden Resultate implizieren zumindest im Falle der Lippenpalten-Chirurgie, dass operative Verfahren, welche die ringförmige Natur des MOO berücksichtigen, Ansätzen vorgezogen werden sollten, die dies nicht tun. Zukünftige FE Simulationen des MOO sollten zudem die umliegenden Muskeln miteinbeziehen, da sie zur Lippenmotilität beitragen. Eine Integration dieser Strukturen in ein simulierbares Gesamtmodell würde einen ganzheitlichen Ansatz repräsentieren, der aber momentan jenseits der Rechenleistungen normaler PCs liegt, besonders dann, wenn Echtzeitfunktionalität gewünscht wird. Virtuelle Chirurgie erlaubt dem Operateur eine Vielzahl an Szenarien für eine bevorstehende Operation gefahrlos durchzuspielen, um so das postoperative Ergebnis bestmöglich vorherzusagen. Eine Simulation eines Rekonstruktionsverfahrens bei Spaltfehlbildungen sollte daher nicht nur auf der Expertise in normaler und pathologischer Anatomie basieren, sondern auch akkurate Repräsentationen der biophysikalischen Eigenschaften der unterschiedlichen Gewebearten sowie virtuelle Agonist-Antagonist Interaktionen miteinschließen. Die Zukunft wird entscheiden, ob die hier präsentierte Simulationskette einen Schritt in die richtige Richtung darstellt.

Literaturverzeichnis

1. Bitter K. Chirurgische Erstbehandlung der Lippen-Kiefer-Gaumen-Spalten im Jahr 2000. *Mund Kiefer Gesichtschir.* 2000;4(7):49–60.
2. Weichert F, Schröder A, Landes C, et al. Netzgenerierung und Finite-Elemente-Simulation muskulärer Strukturen unter Beachtung korrespondierender histologischer Schnittpräparate. In: *Proc BVM*; 2009. p. 192–96.
3. Frey PJ. *Mesh Generation*. Int Soc Technol Educ Publishing; 2008.
4. Sarfraz M. Object recognition using Fourier descriptors: some experiments and observations. *Proc Int Conf Comp Graph Imaging Visual.* 2006; p. 281–86.
5. Crisfield MA. *Nonlinear finite element analysis of solids and structures*. Chichester: Wiley; 1997.

Kalibrierung eines 3D-Ultraschallsystems mit evolutionärer Optimierung

Susanne Winter, Kai Ritschel, Magdalena Broll, Claudia Dekomien

Institut für Neuroinformatik, Ruhr-Universität Bochum
Susanne.Winter@neuroinformatik.rub.de

Kurzfassung. Der Einsatz von 3D-Ultraschall ist für zahlreiche medizinische Applikationen interessant. Ein 3D-Ultraschall Datensatz kann aus 2D-Schichtbildern rekonstruiert werden, wenn deren Positionen im Raum bekannt ist. Mittels eines Trackingsystems kann die Position eines Schallkopfes im Raum verfolgt werden. Die Beziehung von Bildebene und Schallkopf muss durch eine Kalibrierung bestimmt werden. Mit Hilfe einer evolutionären Optimierung, der CMA, wurde eine schnelle und robuste Ultraschallkalibrierung realisiert. Es konnte gezeigt werden, dass der Algorithmus auch bei großer Startabweichung robuste Ergebnisse liefert

1 Einleitung

Ultraschall ist das am häufigsten eingesetzte bildgebende Verfahren in der Medizin. Seit einigen Jahren bietet die 3D-Bildgebung mit Ultraschall zahlreiche neue Möglichkeiten. Hierfür können entweder 3D-Schallköpfe verwendet werden, die direkt einen Volumendatensatz, allerdings mit beschränktem Sichtfeld, akquirieren, oder es werden Volumendatensätze aus zweidimensionalen Schichtbildern rekonstruiert.

Ansätze, die durch alleinige bildbasierte Registrierung Volumendaten erstellen, sind nicht für jede gescannte Region umsetzbar und auf Grund der geringen Genauigkeiten z.B. nicht für die intraoperative Navigation geeignet. Daher werden häufig Methoden genutzt, die eine Kombination von Positionsdaten des Ultraschallkopfes mit Ultraschallbilddaten nutzen. Für die Positionsverfolgung werden herkömmliche Trackingsysteme verwendet, wobei ein Referenzmarker am Schallkopf befestigt wird, dessen Position im Raum gemessen wird. Um die relative Position des Schallkopfes zu den Bilddaten zu bestimmen, wird eine Kalibrierung benötigt. Diese liefert die Koordinatentransformation zwischen Referenzmarker und Bildebene. Es existieren eine Reihe von Ansätzen für die Kalibrierung [1, 2]. Viele dieser Verfahren sind sehr aufwendig und die Genauigkeit ist entweder stark benutzerabhängig, oder erfordert einen speziellen Messaufbau.

Unser Ziel ist es eine einfache und trotzdem genaue Kalibrierungsmethode zu entwickeln, die ohne großen Aufwand von beliebigen Nutzern durchgeführt werden kann. Als Kalibrierphantom haben wir ein einfaches Kugelphantom gewählt [3]. Die Bestimmung der Koordinatentransformation wurde als Optimierungsproblem formuliert, welches mit einem evolutionären Algorithmus, dem Kovarianz

Matrix Adaptation-Verfahren (CMA) [4, 5], gelöst wurde. Dieses Optimierungsverfahren hat sich, insbesondere bei Problemstellungen mit Bezug zu realen, und damit verrauschten, Daten [6, 7], als schnell, zuverlässig und robust erwiesen.

2 Material und Methode

Die Ultraschalldaten wurden mit einem Telemed-System (Echo Blaster 128) bei Verwendung eines 9 MHz *linear-array* Schallwandlers aufgenommen. Die Position des Ultraschallkopfes wurde mit einem optischen Trackingsystem (NDI Polaris) und passiven Referenzmarkern (Abb. 1a)) bestimmt.

Als Kalibrierphantom diente eine mit NaCl-Lösung gefüllte Box, in der ein Tischtennisball befestigt wurde. Um die Bilddaten aus möglichst verschiedenen Schallkopfpositionen akquirieren zu können, wurden zwei Seiten der Box mit einer ultraschalldurchlässige Membran (Abb. 1a)) versehen.

Unter Variation der Aufnahmeposition sowie des Aufnahmewinkels, wurde der Tischtennisball, bei Freihandführung des Schallkopfes, aufgenommen. Eine solche Aufnahme resultierte in ca. 300 Einzelschichten, in denen sich jeweils ein Teil der Kugel als Kreis abbildet (Abb. 1b)).

Mittels Houghtransformation wurden für jede Schichtbildaufnahme der Kreismittelpunkt und der Radius bestimmt. Anhand des maximalen detektierten Radius wurde der Abstand d_i jedes Schichtbildes S_i zum Kugelmittelpunkt bestimmt. Für jedes Schichtbild S_i existieren zwei mögliche Kugelmittelpunkte, k_{i1} und k_{i2} , die jeweils im Abstand d_i auf der Senkrechten durch den Kreismittelpunkt liegen.

Ziel der Kalibrierung ist es, die Transformationsmatrix C zu bestimmen, welche die Beziehung zwischen der Referenzbasis am Schallkopf und der Ultraschallbildebene beschreibt. Die Transformation C ist optimal bestimmt, wenn

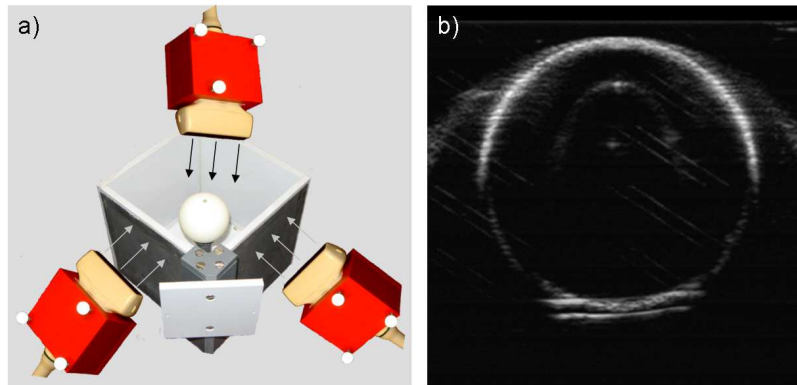


Abb. 1. a) Ultraschallaufnahme für die Kalibrierung; das Kugelphantom wird von 3 Seiten mit einem getrackten Ultraschallwandler aufgenommen b) einzelne Ultraschallschichtaufnahme des Kugelphantoms.

die jeweils richtigen Kugelmittelpunkte nach der folgenden Transformation aufeinander liegen:

$$\hat{k}_{ij} = T_i * C * k_{ij} \text{ mit } j = \{1, 2\} \quad (1)$$

Dabei beschreibt T_i die Positionsmatrix des Schallwandlers im Koordinatensystem des Kugelphantoms, für ein Schichtbild S_i . \hat{k}_{ij} bezeichnet die transformierten möglichen Kugelmittelpunkte.

Da reale, verrauschte Daten die Basis der Kalibrierung bilden, kann die optimale Transformation nicht direkt berechnet werden, sondern wird mittels Optimierungsfunktion bestimmt. Wir haben hierfür einen evolutionären Ansatz, den CMA-Algorithmus, gewählt. Die zu optimierenden Parameter p sind drei Translations- und drei Rotationsparameter. Als Optimierungskriterium dient ein Distanzmaß, welches den quadratischen Abstand der korrekten Kugelmittelpunkte zum Schwerpunkt dieser Kugelmittelpunkte beschreibt. Um dieses zu berechnen, werden zunächst alle Punkte k_{ij} wie oben beschrieben transformiert, nachfolgend wird von jedem Punktpaar derjenige Punkt ausgewählt, der dem Schwerpunkt am nächsten liegt. Aus diesen Punkten wird das Optimierungskriterium bestimmt.

Zur Bestimmung der Reichweite sowie der Präzision des Algorithmus wurden zunächst durch mehrfache lokale Optimierung für jeden Datensatz die optimalen Parameter p_{opt} bestimmt. Ausgehend von diesen Parametern wurden 100 Startpositionen mit großen Abweichungen (bis $\pm 180^\circ$ und ± 1000 mm) generiert. Für jede dieser Startpositionen wurde die Kalibrierung durchgeführt und jeweils eine Transformationsmatrix bestimmt.

Zielpunkte in der Ultraschallbildebene wurden mit der ermittelten Kalibriermatrix und der optimalen Kalibriermatrix C_{opt} transformiert. Der RMS-Fehler, welcher die Abweichung zwischen den transformierten Punkten beschreibt, dient als Präzisionsmaß einer Kalibrierung.

Um die Qualität der Kalibriermatrizen C visuell zu beurteilen wurden aus den akquirierten Datensätzen des Kugelphantoms mittels der Matrizen ein 3D-Datensatz rekonstruiert.

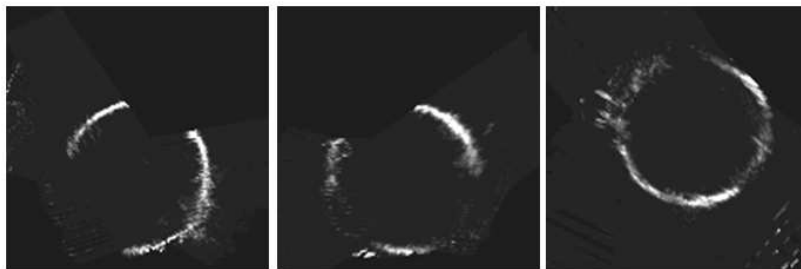


Abb. 2. Rekonstruierter Datensatz; Schnittbilder in drei verschiedenen Raumrichtungen.

3 Ergebnisse

Insgesamt wurden auf zwei Ultraschalldatensätzen jeweils 100 Kalibrierungen durchgeführt. Die Startpositionen wurden um das Optimum des jeweiligen Datensatz gleich verteilt, wobei die Abweichungen bis zu $\pm 180^\circ$ und ± 1000 mm vom Optimum betragen. Auf dem ersten Datensatz ergab sich ein mittlerer RMS-Fehler von 0.28 mm vom bestimmten Optimum, 98 % aller Kalibrierungen hatten eine Abweichung von weniger als 1 mm. Der zweite Datensatz wies einen mittleren RMS-Fehler von 0.65 mm auf, wobei 80 % aller Kalibrierungen unter 1 mm und 87 % aller Kalibrierungen unter 2 mm lagen.

Mit den optimalen Kalibriermatrizen, die aus den beiden Datensätzen berechnet wurden, sind 3D-Volumen rekonstruiert worden. Die Kugeln (Abb. 2) wiesen keine Verzerrungen auf.

4 Diskussion

Der vorgestellte Algorithmus zur 3D-Ultraschall Kalibrierung zeigt Resultate mit guter Präzision, wobei die Optimierung zuverlässig und weitreichend in jedem Fall ein Ergebnis in der Nähe des Referenzoptimums lieferte. Eine robuste Kalibrierung bei Abweichungen von bis zu 180° und 1000 mm vom Optimum hat gezeigt, dass keine besondere Initialisierung der Startparameter notwendig war, um das optimale Ergebnis zu erhalten. Eine mittlere Abweichung von 0.28 mm und 0.65 mm liegt innerhalb der geforderten Genauigkeit von Navigationssystemen.

Im nächsten Schritt soll die Streuung der Kalibrierung bei einer Reihe verschiedener Aufnahmen für verschiedene Schallköpfe evaluiert werden. Des Weiteren ist geplant, die Kalibrierung mit hochpräzisen Kalibrierungen, die mit Laboraufbauten erreicht werden, zu vergleichen.

Literaturverzeichnis

1. Hsu P, Prager R, et al. Freehand 3D ultrasound calibration: a review. *Adv Imaging Biol Med.* 2008;1:47–84.
2. Mercier L, Lango T, et al. A review of calibration techniques for freehand 3D ultrasound systems. *Ultrasound Med Biol.* 2005;31(4):449–71.
3. Brendel B, Winter S, Ermert H. A simple and accurate calibration method for 3D freehand ultrasound. *Biomed Tech.* 2004;49(2):872–73.
4. Hansen N, Kern S. Evaluating the CMA evolution strategy on multimodal test functions. *Proc PPSN.* 2004;3242:282–91.
5. Suttorp T, Hansen N, Igel C. Efficient covariance matrix update for variable metric evolution strategies. *Mach Learn.* 2009;75:167–97.
6. Winter S, Brendel B, et al. Registration of CT and intraoperative 3D ultrasound images of the spine using evolutionary and gradient-based methods. *IEEE Trans Evol Comput.* 2008;12(3):284–96.
7. Winter S, Brendel B, Igel C. Registration of bone structures in 3D ultrasound and CT data: comparison of different optimization strategies. In: *Proc CARS.* 2005;1281:242–47.

Non-Stationary CT Image Noise Spectrum Analysis

Michael Balda¹, Björn J. Heismann^{1,2}, Joachim Hornegger¹

¹Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen

²Siemens Healthcare, Erlangen

`michael.balda@informatik.uni-erlangen.de`

Abstract. We investigate the spatial dependency of noise characteristics in CT images. The perceived image quality depends on the noise-granularity. Especially in low-dose applications the noise granularity influences the diagnostic value. A model is presented, that provides two-dimensional, stationary noise realizations for arbitrary image pixel locations from which two-dimensional Noise Power Spectrum estimates can be computed. It fully incorporates the CT reconstruction process for (indirect) fan-beam reconstruction, the quarter offset of the detector channels and the detector noise characteristics. It can be used with simulated and measured data and allows for the assessment of spectral noise characteristics for arbitrary objects.

1 Introduction

The noise in the reconstructed CT image is influenced by various parameters. When planning components of a Computed Tomography scanner, like detectors or X-ray tubes, it is necessary to tune their parameters in a way that leads to the best achievable objective image quality. In terms of software aspects, different reconstruction techniques and several noise-reduction algorithms like sinogram pre-processing exist, which have to be evaluated not only in terms of total noise reduction but also on their influence on the image quality. These tasks require methods to assess the image noise characteristics.

Fig. 1 illustrates the typical noise characteristics of a CT image. The noise texture shows an isotropic grain structure at the center. Towards the boundary it decreases and becomes increasingly oriented. This effect can be observed in both images of different water phantoms. There are models which analytically compute the noise variance propagation for the inverse Radon transform [1], more precise noise propagation models also include local correlations introduced by various reconstruction steps [2]. Additionally, several approaches have been made, which model the noise propagation from the detector noise transfer function to the image noise power spectrum (NPS) [3], [4] but these models are limited by the fact, that the assumptions on the noise transfer functions (like detector- or focus-noise transfer function) are only valid for the image center. Objective quality measures of a CT image, however, rely on an assessment of the noise characteristics throughout the whole image.

2 Materials and Methods

We develop a model to generate two-dimensional realizations of the image noise and NPS estimates for arbitrary image positions. It uses a projection noise model. For any pixel position and projection angle, an ensemble of projection noise values at the associated detector position is generated. This local detector noise signal then undergoes all steps of the reconstruction algorithm that influence the noise characteristics. Back-projection of the according noise signals for each angle yields a stationary, two-dimensional noise realization for this image location. This model gives an insight into the noise shaping properties of the filtered back-projection (FBP). Unlike analytical noise propagation methods, it can be easily adapted to any FBP-based reconstruction algorithm and various effects can be incorporated into the model without greatly increasing its overall complexity.

The characteristics of this noise realization will differ throughout the image and resemble the CT image noise at a selected location $\mathbf{p} = (p_x, p_y)^T$. The value of an image pixel depends on the attenuation of all beams passing through this pixel. These beams have associated projection angles $\phi \in [0; \pi)$. The model creates a stationary local noise image patch $i_{\mathbf{p}}(\mathbf{x})$ using the following steps for all projection angles ϕ ([5] for a concise description of the FBP):

1. *Find measured attenuation value:* The detector coordinate $\gamma(\mathbf{p}, \phi)$ can be computed as follows:

$$\gamma(\mathbf{p}, \phi) = \left(\text{atan2} \left(\frac{r(\mathbf{p}) \cdot \cos(\phi - \phi_T(\mathbf{p}))}{d_{\text{SOC}} + r(\mathbf{p}) \cdot \sin(\phi - \phi_T(\mathbf{p}))} \right) + \frac{\beta}{2} - \beta_a \right) \cdot \frac{1}{\Delta\beta} \quad (1)$$

The complete fan angle is β , $\Delta\beta$ is the fan angle between two neighboring detector channels, β_a is the alignment angle due to the quarter offset ($\beta_a = 0$ for no quarter offset). The distance between X-ray source and origin is denoted d_{SOC} , the polar coordinates of \mathbf{p} are denoted $r(\mathbf{p})$ for the distance to

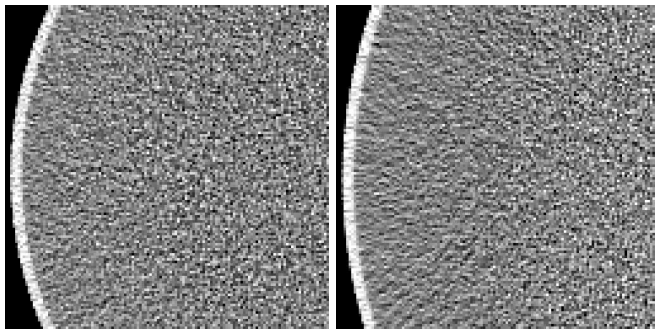


Fig. 1. Excerpts (125 mm×125 mm) of CT images of water phantoms: 300 mm diameter (left), 400 mm diameter (right).

the origin and $\phi_T(\mathbf{p})$ for the polar angle. Using $\gamma(\mathbf{p}, \phi)$, the according attenuation value $\mu(\gamma(\mathbf{p}, \phi), \phi)$ can be interpolated from the measured projection data.

2. *Initialize projection noise model:* The projection noise model provides a noise distribution for the attenuation values $\mu(\gamma(\mathbf{p}, \phi), \phi)$. The noise in the attenuation values measured by the detector consists of electronics noise, quantization noise and quantum noise. The quantum noise depends on the number of incoming X-ray quanta. A detailed description on how quantum noise affects the detected attenuation values can be found in [1]. In practice, the noise effects can either be modeled or assessed by calibration measurements. Usually, the detector noise signal can be assumed to be uncorrelated and white. In practice minor correlations between the detector channels are introduced by optical cross-talk.
3. *Generate virtual detector noise signals:* A virtual, stationary detector noise signal is generated, that has constant noise distribution parameters throughout all channels. In case of rebinning to parallel beam geometry, an additional reading is generated with identical parameters. It represents the projection from the opposite direction. If the detector noise cross-talk behavior should be considered, these signals have to be convoluted with the impulse response of the detector.
4. *Rebin virtual noise signal:* The rebinning also influences the NPS. It consists of two steps. First the beams are reordered to parallel beam geometry with inhomogeneous ray distances. This requires a 2-D interpolation in the sinogram. In case of a quarter offset, the beams from opposite projection directions are then sorted into one single reading with the double amount of channels, finally the readings are resampled to homogeneous ray distances which requires a 1-D interpolation. In case of direct FBP, this step is omitted.
5. *Filtered back-projection:* The rebinned virtual noise signal is filtered and back-projected. The resulting noise pattern is stationary and resembles noise characteristics as this specific location in the image.

2.1 Algorithm

The algorithm can be summarized as follows.

1. Set projection angle $\phi = -\pi/2 + \Delta\phi \cdot r$ and $\Delta\phi = \pi/n_r$.
2. Calculate $\gamma(\mathbf{x}, \phi)$ according to (1) and interpolate reading $d_\phi[c]$ at γ to get attenuation value $\mu(\phi)$.
3. Initialize detector noise generator with $\mu(\phi)$.
4. Generate two vectors of noise signal realizations $\tilde{\mathbf{d}}_{n,i}[c]$ with $c \in [0; N_d]$ and $i \in \{1, 2\}$.
5. Perform rebinning on $\tilde{\mathbf{d}}_{n,1}$ and $\tilde{\mathbf{d}}_{n,2}[c]$ to get $\hat{\mathbf{d}}_n[c_p]$.
6. Back-project $\hat{\mathbf{d}}_t[c_p]$ onto the noise patch $i_{\mathbf{x}}$.
7. if $r < N_r$: Set $r = r + 1$ and GOTO 2;

The length of the detector noise signal N_d has to be chosen according to the desired size and field of view (FOV) of the noise resulting image patch $i_{\mathbf{x}}$ so that the image patch is fully covered by the back-projected values of $\tilde{\mathbf{d}}_t[c]$, n_r is the number of desired rotation angle samples.

3 Results and Discussion

The algorithm was tested with several measured and simulated phantoms. Fig. 2(a) shows the reconstruction of a simulated phantom with marked voxel locations. This phantom is especially suited to demonstrate the properties of the proposed method, since the two strongly attenuating bone structures have a dominating influence on the noise structure of the whole image. Fig. 3 displays the respective stationary noise patches and the corresponding NPS estimates. Dominating noise components in directions of strong attenuation (between the two bones) can be observed. This leads to an increasingly directed noise grain structure. For the pixel close to the border, there is no such dominant influence on the noise direction, this leads to a more homogeneous noise structure. The overall noise is stronger for pixels close to the center and / or near strong attenuators. For testing the noise model on measured data, an elliptic water phantom of 450 mm width and 225 mm height was used. Fig. 2(b) shows the measured noise standard deviation from a circular ROI of radius 50 mm shifted along the longer axis of the ellipse. The corresponding model estimates are generated with the introduced method, an analytic forward projector and a calibrated noise model. The measurement and the values computed with the model on simulated data show a good agreement with relative deviations below 5%.

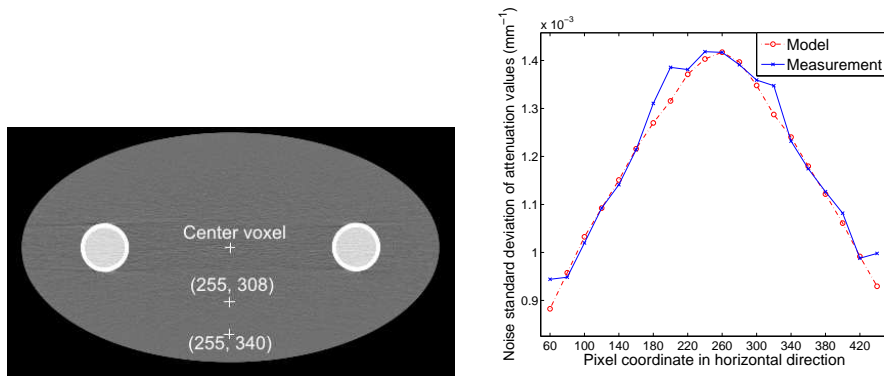
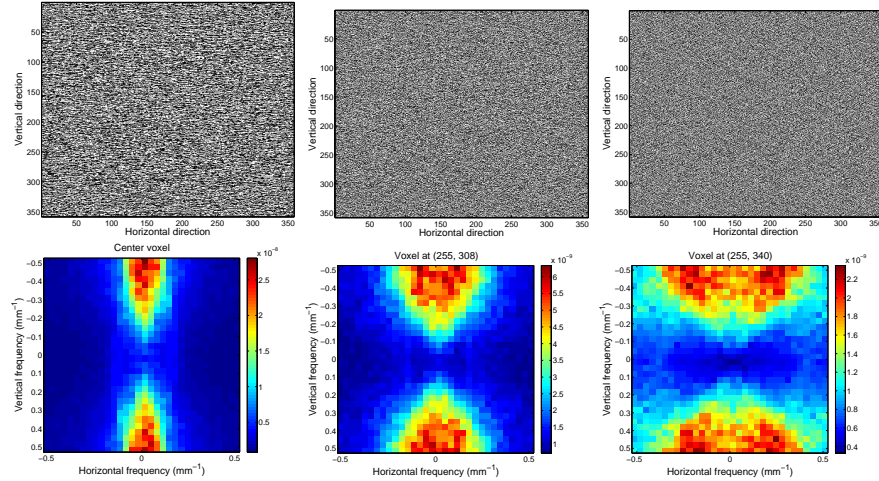


Fig. 2. Left: Reconstruction of an elliptical water phantom with two femur bone structures. Noise patches are provided in Fig. 3 for marked voxel locations; Right: Comparison of noise standard deviation in from measured data and model.

Fig. 3. Top row: Noise patches for voxels marked in Fig. 2(a) (window center: 0, width: 4.2 HU); bottom row: normalized NPS estimates computed from noise patches above.



4 Conclusion

We have introduced a method to generate stationary noise patches and Noise Power Spectrum estimates that resemble the noise characteristics of arbitrary voxel positions of CT images. The NPS estimates show the influence of object structures and reconstruction properties on the image noise structures. The model can be easily adapted to various FBP-based reconstruction techniques. This method can be used as a valuable tool to evaluate the influence of reconstruction parameters such as reconstruction kernels or noise adapted sinogram filters on the quality of the reconstructed image.

References

1. Buzug TM. Einführung in die Computertomographie: Mathematisch-physikalische Grundlagen der Bildrekonstruktion. Berlin: Springer; 2005.
2. Borsdorf A, Kappler S, Raupach R, et al. Analytic noise propagation for anisotropic denoising of CT images. In: IEEE Nucl Sci Symp Conf Rec; 2008. p. 5335–38.
3. Riederer SJ, Pelc NJ, Chesler DA. The noise power spectrum in computed X-ray tomography. *Phys Med Biol.* 1978;23(3):446–54.
4. Kijewski MF, Judy PF. The noise power spectrum of CT images. *Phys Med Biol.* 1987;32(5):565–75.
5. Kak AC, Slaney M. Principles of Computerized Tomographic Imaging. Soc Ind Appl Math Publishing; 2001.

Automatisierte quantitative Analyse der Zellzusammensetzung von bronchoalveolaren Spülungen

Julian Wörmann¹, Andrea Braun², Martin Mempel^{2,3}, Karl-Hans Englmeier¹,
Peter Hamm¹

¹Institute for Biological and Medical Imaging, Helmholtz Zentrum München

²Division of Environmental Dermatology and Allergy

Helmholtz Zentrum München/Technische Universität München (TUM)

ZAUM - Center for Allergy and Environment (TUM), Germany

³Department of Dermatology and Allergy Biederstein, TUM, Germany

julian.woermann@helmholtz-muenchen.de

Kurzfassung. Eine in der Immunologie häufig verwendete Methode zur Bewertung von Entzündungsreaktionen ist die quantitative Auswertung von Leukozyten, wie z.B. die Bestimmung der Zusammensetzung der Entzündungszellen in bronchoalveolaren Spülungen (BAL) bei Lungenerkrankungen. Die vorliegende Arbeit stellt eine bildbasierte Applikation zur automatisierten BAL-Zellklassifikation in Mikroskopaufnahmen vor, welche gegenüber der gängigen manuellen Auswertung sowohl einen erheblichen Zeitgewinn, als auch einen Fortschritt bezüglich der Klassifikationsgenauigkeit darstellt. Die automatische differenzierte Zellzählung erfolgt in vier Schritten und erlaubt durch die Einführung eines Konfidenzparameters eine interaktive Validierung des Anwenders, wodurch eine sehr hohe Sensitivität erzielt wird.

1 Einleitung

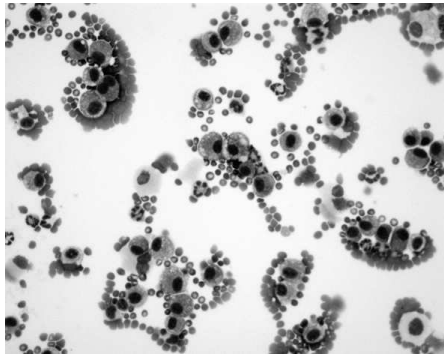
In der Allergieforschung wird anhand von Mausmodellen die allergische Reaktion der Lunge anhand der Zellzusammensetzung der bronchoalveolaren Spülung (BAL) bestimmt. Die durch Spülung der Lunge gewonnenen Cytospin-Präparate von BAL Zellen werden durch eine differenzierte Zellfärbung gefärbt und anhand ihrer morphologischen Besonderheiten charakterisiert. Dabei ist die absolute und relative Häufigkeit von insgesamt vier verschiedenen Leukozyten von Interesse: Makrophagen, Lymphozyten, Eosinophile und Neutrophile. Die quantitative Auswertung ist ein wichtiger Parameter in der Bestimmung und Bewertung der Stärke verschiedener Atemwegserkrankungen (z.B. Asthma). Der hohe Zeitaufwand der manuellen Zellzählung durch einen Experten begründet sich vor allem darin, dass für eine statistische Auswertung eine hinreichende Anzahl an Proben ausgezählt werden muss. Des Weiteren erschweren geringe Vergrößerungen dem Laboranten die Auszählung der Zellen am Mikroskop, was zu einem erhöhten Fehlerfaktor und somit zu einer verringerten Sensitivität führt.

Bestehende Ansätze um die bildbasierte differenzierte Zellzählung zu automatisieren, führen eine Segmentierung der Zellkerne und Zellplasmen via Watershed-Transformation [1], Thresholding [2] oder einer Kombination aus Thresholding und Fast-Marching [3] durch. Die speziell für die Analyse der BAL Zellen vorgestellte Methode von Lezoray und Lecluse [4] verbindet Histogramm-Clustering mit anschließender Color-Watershed-Transformation, um die Zellen zu segmentieren, und verwendet entsprechend der Färbung Farbmerkmale im HSV-Raum und Texturmerkmale der Nuklei und Plasmen. Die genannten Methoden sind jedoch nur teilweise auf dem vorliegenden Falle der Differenzierung von *einzelnen* Leukozytentypen anwendbar, für die nach unserem Wissen noch keine vorgestellte automatisierte Lösung existiert.

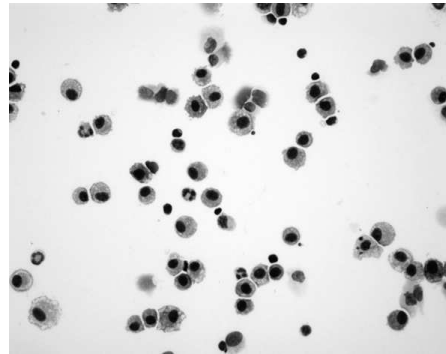
2 Material und Methoden

Die durch Cytozentrifugation auf einem Objektträger fixierten BAL Zellen werden durch "Diff Quick Staining Set" (Medion Diagnostic) mit Eosin und Thiazin gefärbt. Die Differenzierung der Zelltypen erfolgt anhand von morphologischen Kriterien: Eosinophile können anhand ihres polymorphoiden Kerns und des kationischen Zytoplasmas (rote Granula) erkannt werden; das Zytoplasma der Neutrophilen hingegen wird nicht gefärbt, ihr Zellkern besteht meistens aus 3-4 Segmenten; Makrophagen können von Lymphozyten anhand ihres höheren Verhältnis an Zytoplasma zu Kern und den weniger kondensierten Kern abgegrenzt werden.

Das zur Aufnahme der Bilder verwendete Mikroskop ist ein Leica DM LB mit einer Leica DC 300F Kamera, welche RGB-Bilder mit einer Auflösung von 1030 x 1300 Pixel erzeugt. Ausgehend von der digitalen Erfassung der eingefärbten Proben (Abb. 1), erfolgt der Ablauf der Applikation in vier Schritten: Vor-



(g) Beispiel einer schwierig zu analysierenden Probe, aufgrund des hohen Anteils an Erythrozyten



(h) Beispiel einer einfach zu analysierenden Probe

Abb. 1. Gegenüberstellung von schwierig und einfach zu analysierenden Cytospin-Präparaten der BAL Zellen.

verarbeitung, Segmentierung, Merkmalsextraktion und Klassifizierung. In der ersten vorverarbeitenden Stufe wird Bildrauschen mit Hilfe eines nicht-linearen und kantenerhaltenden Filters [5] entfernt und die inhomogene Bildausleuchtung per homomorphischer Filterung beseitigt. Besonderes Augenmerk gilt dem zweiten Schritt, der Segmentierung von Zellkernen und Zellplasma (Abb. 2), deren Ergebnis unmittelbare Auswirkung auf die Klassifizierung hat. Als Grundlage hierfür findet eine marker-basierte Watershed-Transformation Anwendung. Die regionalen Maxima einer Distanztransformation, die auch bei sich berührenden Zellrändern das Zentrum einer Zelle definieren, bilden die gesuchten Marker.

Die Segmentierung der Zellkerne erfolgt auf Basis von Schwellwertoperationen im HSV-Farbraum und einer nachgestellten Anwendung morphologischer Operatoren. Abbildung 2 zeigt die mittels Schwellwert extrahierten Zellkerne und die durch die Watershed-Transformation erhaltenen Zellkonturen im Vergleich zum Originalbild. Da Makrophagen und Lymphozyten i.Allg. nur einen Zellkern besitzen, werden die Regionen dem zuvor erstellten Markerbild hinzugefügt, was die Segmentierung von Zellclustern verbessert. Die im dritten Schritt vorgenommene Extraktion von diskriminierenden Merkmalen basiert auf den Farb- und Formmerkmalen der vier Zelltypen. Dazu zählen die Größe der Zelle, die Größe der Zellkerne und deren Anzahl, das Verhältnis zwischen Zellkern und Zytoplasma, Farbton und Sättigung von Kern und Plasma und formbasierte Merkmale wie die Rundheit und die Kompaktheit der Zellkerne.

Vor der eigentlichen Klassifizierung wird mittels eines Entscheidungsbaumes jede Region aufgrund ihrer extrahierten Merkmale einer von vier Klassen zugeordnet. Zunächst erfolgt die Trennung der Regionen von allen Objekten, die nur einen einzigen, kompakten Zellkern aufweisen. Makrophagen und Lymphozyten erfüllen dieses Kriterium und werden weiterhin aufgrund ihrer unterschiedlichen Zellgröße voneinander getrennt. Aus den übrigen Regionen werden die Segmente extrahiert, die in ihrem Schwerpunkt einen höheren Grauwert aufweisen, als an ihrem äusseren Rand. Diese Vorgehensweise begründet sich darin, dass die in

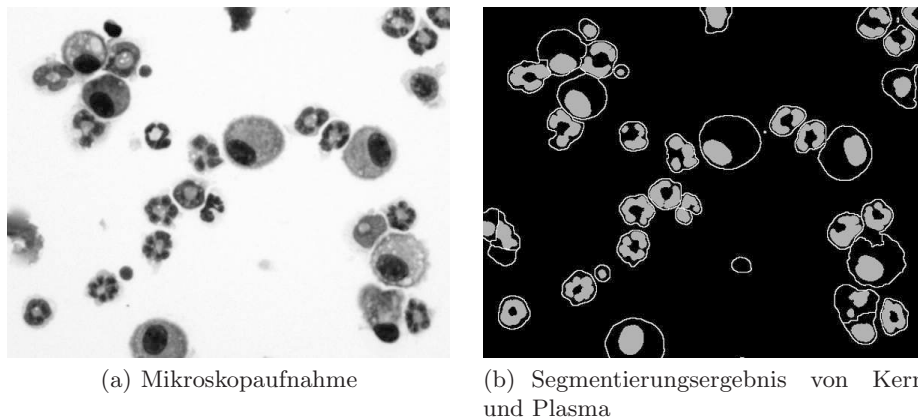


Abb. 2. Segmentierungsergebnis eines Ausschnitts einer zu klassifizierenden Zellprobe.

Tabelle 1. Klassifikationsergebnis der vorgestellten Applikation anhand 18 repräsentativer Beispiele von Cytospin-Präparaten der BAL Zellen.

(a) Richtig positiv Rate bezogen auf die Expertenanzählung			(b) Klassifikationsergebnis bezogen auf alle segmentierten Objekte		
	Soll	rp		Anzahl	Prozent
Makrophagen	470	373	Objekte	1450	100 %
Lymphozyten	145	97	davon falsch	97	6,69 %
Eosinophile	466	339	klassifiziert		
Neutrophile	34	20			
			zur Validierung	524	36,14 %
Summe	1115	829	davon Zellschrott	317	60,50 %
Prozent	100 %	74,35 %	davon Zellen	207	39,50 %

der Regel am Rand befindlichen Zellkerne der Eosinophile und Neutrophile eine höhere Sättigung besitzen, als das Zytoplasma im Zentrum. Dies bewirkt eine Trennung der Zellen von Farbresten und Zellbruchstücken. Letztere Regionen, die keiner der drei Gruppen zugeordnet werden konnten, gehen somit nicht in die Klassifikation mit ein und stehen dem Benutzer zur Validierung zur Verfügung.

Die Klassifikation erfolgt nach der Methode des geringsten Abstandes zu Prototypen der vier Objektklassen, die zuvor aus einer Mittelwertbildung der Merkmale von jeweils zehn repräsentativen Zellen auf Trainingsbildern berechnet wurden. Die Klassifizierung verläuft dabei separat in den zuvor erstellten Gruppen, wobei die ersten zwei Gruppen jeweils einen Prototypen besitzen. Die segmentierten Regionen werden der Klasse zugeordnet, wenn ihre Distanz einen festgelegten Schwellwert (Konfidenzparameter) nicht überschreitet. Die Regionen der dritten Gruppe (Eosinophile und Neutrophile) werden zusätzlich einer Distanzmessung zu den Klassenprototypen unterzogen, um die jeweilige Klassenzugehörigkeit festzustellen. Der Konfidenzparameter gibt Aufschluss darüber, ob eine Zelle mit Sicherheit einer der Klassen zugeordnet wird, oder ob die Region als nicht eindeutig klassifiziert markiert und dem Anwender zur interaktiven Validierung hervorgehoben wird.

3 Ergebnisse

Die Ergebnisse der vorgestellten Applikation wurden anhand 18 repräsentativer Beispiele von Cytospin-Präparaten der BAL Zellen ermittelt, die neben den gesuchten Leukozyten auch Erythrozyten, Zellschrott und Farbrückstände enthalten. Abbildung 1 verdeutlicht die Variation der Zellzusammensetzung in verschiedenen Proben. Tabelle 1(a) zeigt die von der Applikation richtig klassifizierten Zellen in Bezug auf die manuelle Auswertung von einem Experten. Die Klassifizierung erreichte dabei eine Rate von insgesamt 74,35 % richtiger Zuordnung

zu einer der vier Leukozytentypen. Tab. 1(b) verdeutlicht die geringe falsch-positiv Rate. Bezogen auf alle zu klassifizierenden Objekte, wurden 6,69 % dabei einer falschen Klasse zugeordnet und 36,14 % als nicht eindeutig klassifizierbar markiert und dem Anwender zur Validierung weitergegeben. Dabei befanden sich unter den zu validierenden Objekten 60,50 % richtig erkannte Erythrozyten, Farbrückstände und Zellschrott, und 39,50 % Zellen, die entweder zuvor aussortiert, oder keinem Prototypen zugeordnet werden konnten.

Die Implementierung wurde in MatLab auf einem Intel Core 2 Quad 2,66 GHz mit 4 GByte RAM realisiert und benötigte für die Segmentierung durchschnittlich 93,1 Sekunden und für die Klassifikation 30,8 Sekunden Rechenzeit.

4 Diskussion

Das Hauptaugenmerk der vorgestellten Methode liegt vorrangig auf einer Segmentierung aller in der Probe enthaltenen Objekte. Die erreichte Sensitivität von 0,929 verdeutlicht, dass annähernd alle zu analysierenden Objekte segmentiert wurden. Die anschließende Klassifizierung zeigt gute Ergebnisse, vor allem in Hinblick auf die geringe falsch-positiv Rate von 6,69%.

Die Ursachen von Fehlklassifikationen liegen zum einen in der stark schwankenden äußerlichen Form von Zellen gleichen Zelltyps. Zum anderen beeinträchtigen unzureichend segmentierte Zellen oder Zellkerne die vor der Klassifikation stattfindende Einteilung der Probe in einer der Gruppen. Des Weiteren kann der Farbton und die Sättigung der Kerne sowie der Plasmen einerseits durch die Färbung, und andererseits durch ein verändertes Mikroskop-Setup stark variieren, was ebenfalls eine eindeutige Zuordnung erschwert.

Ziel der zukünftigen Arbeit ist daher die Anpassung des Algorithmus an sich ändernde Aufnahmebedingungen, sodass der Einfluss von variierenden Aufnahmeparametern wie Farbton, Helligkeit und Sättigung minimiert wird.

Die im Rahmen dieser Arbeit entstandenen Methoden stellen eine bildbasierte Applikation zur automatisierten BAL-Zellklassifikation dar, welche gegenüber der bisherigen manuellen Auswertung sowohl einen erheblichen Zeitgewinn, als auch eine reproduzierbare, präzise und objektivierbare Klassifikation durchführt.

Literaturverzeichnis

1. Dorini LB, Minetto R, Leite NJ. White blood cell segmentation using morphological operators and scale-space analysis. *SIBGRAPI*. 2007; p. 294–304.
2. Mirčić S, Jorgovanović N. Automatic classification of leukocytes. *J Automat Contr*. 2006;16:29–32.
3. Rehn T, Zerfaß T, Wittenberg T. Berandungsgenaue Segmentierung von Plasma und Nucleus bei Leukozyten. *Proc BVM*. 2007; p. 252–6.
4. Lezoray O, Lecluse M. Automatic segmentation and classification of cells from broncho alveolar lavage. *Image Anal Stereol*. 2007;26:111–9.
5. Smith SM, Brady JM. SUSAN - A new approach to low level image processing. *Int J Computer Vis*. 1997;23(1):45–78.

Simultaneous PET and MR Imaging with a Newly Developed 3TMR-BrainPET Scanner

Christoph Weirich, Jürgen Scheins, Michaela Gaens, Lutz Tellmann,
Elena Rota Kops, Joachim Kaffanke, Jon Shah, Hans Herzog

Institute of Neuroscience and Medicine – 4, Forschungszentrum Jülich
c.weirich@fz-juelich.de

Abstract. A prototype of a new bimodal scanner was installed in our laboratory. This scanner combines magnetic resonance imaging (MRI) and positron emission tomography (PET) for brain studies. As the PET detector is located within the bore of the MRI scanner, simultaneous measurements become possible. The MR-component consists of a commercial 3T MRI scanner MAGNETOM Tim-Trio, whereas the PET detector BrainPET has been newly developed. The readout electronics of the PET are based on avalanche photodiodes (APDs) which can be applied in the magnetic field. The inner diameter of the PET ring, which consists of 32 cassettes, is 36 cm so that an MRI head coil can be inserted into this ring. First tests were carried out to analyse the count-rate behaviour and the image resolution. To assess the image quality of simultaneous MR-PET, phantom studies were performed. The image resolution in the centre was between 2.1 and 2.5 mm in x,y and z directions. First MR-PET studies showed no visible artefacts of the MR image and FDG-PET images provided high-resolution delineation of the cortex.

1 Introduction

Combining imaging technologies that integrate the strengths of two modalities, and at the same time eliminate one or more weaknesses of an individual modality, offers the prospect of improved diagnostics, therapeutic monitoring and preclinical research. Research groups have been developing MR-PET scanners for small animal imaging [1] such that anatomical information can be obtained by MRI and at the same time PET can observe metabolic function. Recently, some prototypes of an MR-BrainPET scanner for simultaneous human brain studies were built. The feasibility of MR-PET imaging was demonstrated in a human volunteer [2]. One of these prototypes developed by Siemens and combining an industrial 3T MRI scanner with a newly developed PET detector has recently been installed in our PET laboratory and is currently being tested. Here we report our experience with the new hybrid scanner.

2 Material and Methods

2.1 Design Characteristics

The MRI component of the MR-BrainPET consists of the slightly modified MAGNETOM Tim-Trio with a magnetic field of 3 T (Fig. 1(a)). One modification is the replacement of the standard head coil by a combined transmit/receive coil which causes less attenuation of radiation. The BrainPET component has an outer diameter of 60 cm and fits in the MR bore. It consists of 32 copper shielded cassettes each with six compact detector modules (Fig. 1(b)). The front-end of the detector module has 12×12 LSO crystals of $2.5 \times 2.5 \times 20$ mm³ each which are read out by an array of 3×3 avalanche photodiodes (APDs).

While typical PET scanners use photomultipliers as readout electronics, the presence of a strong magnetic field requires solid state components such as APDs which are not magneto-sensitive.

2.2 Phantom Studies

The countrate performance was tested with a cylinder of 25 cm length and 22.5 cm diameter filled with ¹¹C-solution. The activity at the beginning was 417 MBq. During the measurement of 11 half life times ($T_{1/2}=20$ min) the radioactivity concentration dropped from 80 to 0.3 kBq/cc. The transaxial resolution was studied with a ¹⁸F-filled 0.5 mm thin line source and the axial resolution with a 0.25 mm ²²Na point source. Point spread images were obtained by iterative 3D-OSEM reconstruction [3] using 16 subsets and 6 iterations. Image resolution was determined as the full width at half maximum (FWHM) of a Gaussian fit to a profile, which crossed the maximum voxel of the reconstructed point. To study the image quality systematically and to localise image artefacts, several phantom studies were performed. Correction procedures such as

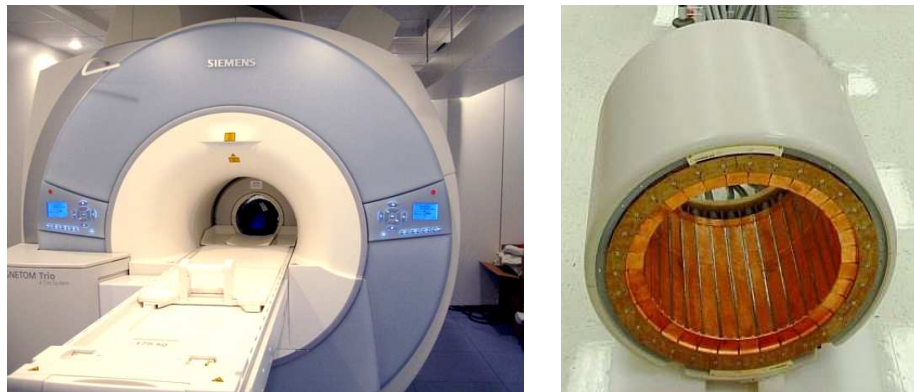


Fig. 1. The 3TMR-BrainPET hybrid scanner. Left: 3T MR scanner MAGNETOM Tim-Trio with BrainPET insert within the MR bore; Right: MR-compatible BrainPET scanner with 32 copper shielded cassettes.

normalisation and corrections for attenuation, Compton scatter and random coincidences have been developed, optimised and tested. These promising results allowed us to perform first qualitative human studies.

2.3 Human Studies

^{18}F -FET-Study: a first combined human MR-PET study was performed in a patient with a malignant brain tumour. After injection of about 200 MBq ^{18}F -fluoroethyltyrosine (^{18}F -FET) [4], the clinical study was performed in the Siemens ECAT Exact HR+ for 50 min. Subsequently the patient volunteered to be scanned in the MR-BrainPET for another 30 min without further radiotracer injection (Fig. 2). During the PET scan different MR sequences were performed.

^{18}F -FDG-Study: another combined human MR-PET study with the radiotracer ^{18}F fluorodeoxyglucose (^{18}F -FDG) was performed to visualise the glucose metabolism of the brain. A tumour patient (non-brain tumour) was injected

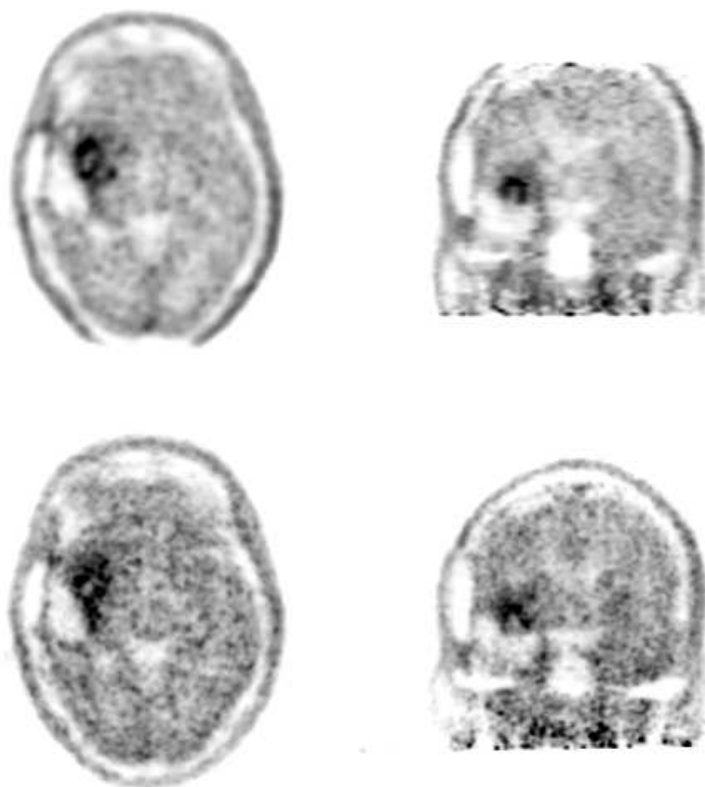


Fig. 2. Upper row: ^{18}F -FET image acquired with the whole body PET scanner HR+. Lower row: ^{18}F -FET image acquired with the BrainPET insert.

342 MBq of ^{18}F -FDG. Prior to the diagnostic whole body scan in the HR+-scanner, the patient volunteered to be scanned in the 3TMR-PET. The acquisition started 30 min after injection (p.i) and was acquired for further 20 min. During the PET scan several MR sequences were performed.

The MR data were processed with the common Syngo software. PET images were reconstructed using the 3D-OSEM reconstruction with 1 subset and 42 iterations (Fig. 3). Prior to the reconstruction varying crystal efficiencies were corrected by a component based normalisation. Since attenuation correction cannot be based on transmission scans due to the lack of a transmission source, it is achieved using attenuation templates derived from transmission measurements in the HR+ and adapted to the individual T1-weighted MP-RAGE image [5].

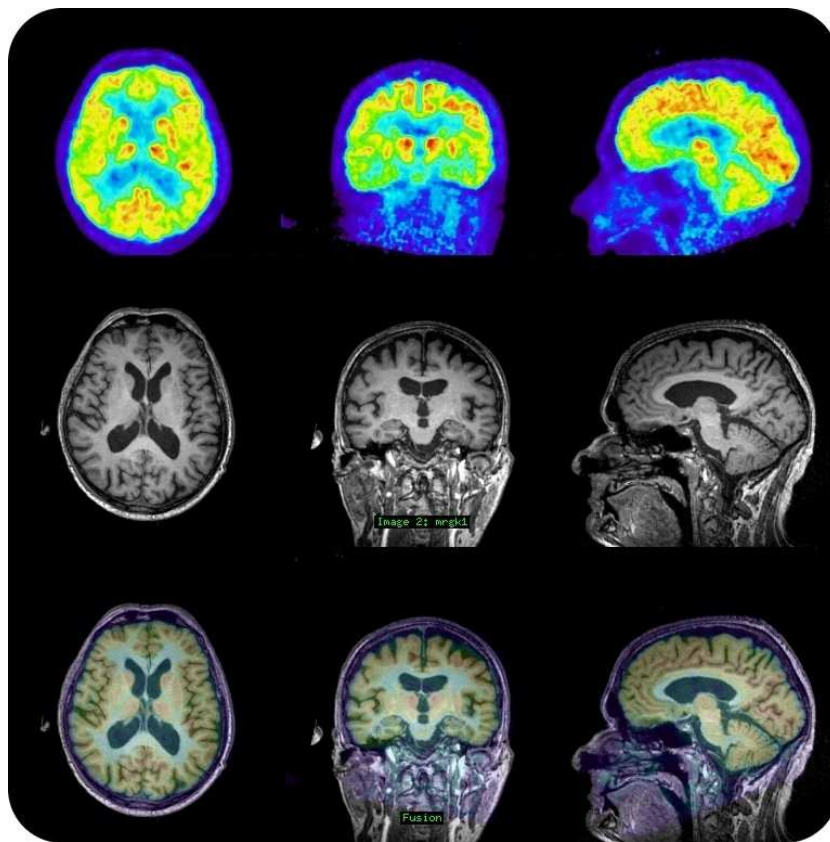


Fig. 3. Simultaneously measured PET and MR images. First row: ^{18}F -FDG image of a 20 min scan. Second row: T1-weighted MR image (MP-RAGE). Third row: Fusion of FDG and MP-RAGE image. Left transversal, middle coronal, right sagittal view. The PET-images were reconstructed using the image reconstruction platform PRESTO [6].

3 Results

Performing the decay experiment with the cylinder, the prompt countrate started to saturate at an activity concentration of about 62 kBq/cc. The netTrue counts as the difference of prompt and delayed counts had a peak of 670000 cps at 52 kBq/cc.

The analysis of the reconstructed point spread images yielded a tangential resolution (FWHM) of 2.4 ± 0.2 mm ($r = 0, 2.5$ and 5 cm) in a central transversel plane and a radial resolution ranging from 2.1 mm via 3.3 mm to 5.5 mm for the three radii. Within the 15 central image planes the axial resolution was 2.5 ± 0.2 mm at $r = 0$ cm and 3.1 ± 0.2 mm at $r = 5$ cm. In comparison to the HR+ scan, the PET images obtained with the BrainPET qualitatively show a higher spatial resolution (Fig. 2). The high resolution of the BrainPET scanner is confirmed by the FDG image of the combined human MR-PET scan in Fig. 3.

4 Discussion

The preliminary tests reported here confirm that simultaneous MR-PET imaging is possible with the newly developed hybrid scanner. In the human studies, MR images show minimal artefacts. The BrainPET detector provides an excellent image resolution. The PET system saturates at activity concentrations which are higher than those expected in brain studies of glucose or receptor metabolism. To date the system is not yet quantitative, since reconstructed images deliver counts instead of activity concentrations.

It is planned to continue with further development of the correction procedures which allow for quantitative PET imaging. Additionally more detailed tests will be performed to obtain further experience in human brain studies.

References

1. Wehrl HF, Judenhofer MS, Pichler BJ, et al. Pre-clinical PET/MR: technological advances and new perspectives in biomedical research. *Eur J Nucl Med Mol Imag.* 2009;36, Suppl 1:56–68.
2. Schlemmer HP, Pichler B, Ladebeck R, et al. Simultaneous MR/PET imaging of the human brain: feasibility study. *Radiology.* 2008;248:1028–35.
3. Lonneux M, Borbath I, Michel C, et al. Attenuation correction in whole-body FDG oncological studies: the role of statistical reconstruction. *Eur J Nucl Med.* 1999; p. 591–8.
4. Pauleit D, Floeth F, Langen KJ, et al. O-(2-[18F]fluoroethyl)-L-tyrosine PET combined with MRI improves the diagnostic assessment of cerebral gliomas. *Brain.* 2005;128:678–87.
5. Rota Kops E, Qin P, Herzog H, et al. Attenuation correction of PET scanning based on MR images. *IEEE Nucl Sci Symp and Med Imag Conf Rec.* 2006.
6. Scheins J, Boschen F, Herzog H. Analytical calculation of volumes-of-intersection for iterative, fully 3-d PET reconstruction. *IEEE Trans Med Imaging.* 2006;25.

Echtzeitfähige Extraktion scharfer Standbilder in der Video-Koloskopie

Sebastian Gross^{1,2}, Martin Schink¹, Thomas Stehle¹, Alexander Behrens¹,
Jens Tischendorf², Christian Trautwein² und Til Aach¹

¹Lehrstuhl für Bildverarbeitung, RWTH Aachen University

²Medizinische Klinik III, Universitätsklinikum Aachen

sebastian.gross@lfb.rwth-aachen.de

Kurzfassung. Die endoskopische Untersuchung des Darms (Koloskopie) ist ein wichtiges Verfahren zur Diagnose und Therapie vieler Erkrankungen. Während der Koloskopie werden u. a. zu dokumentarischen Zwecken oder als Grundlage für die computergestützte Diagnose und Therapie Standaufnahmen erzeugt. Aufgrund der Peristaltik des Darms und der eingeschränkten Kontrolle über das Koloskop ist es schwierig, Standbilder aufzunehmen, die keine Bewegungs- oder Focus-bedingte Unschärfe aufweisen. Eine Verbesserung der Situation kann erreicht werden, indem statt einer einzelnen Aufnahme eine Sequenz von Bildern gewonnen wird, aus denen in Echtzeit mit Hilfe von schnellen Schärfemaßen zur Bewertung der Bildschärfe das detailreichste Bild gewählt wird. Wir stellen drei Schärfemaße vor und vergleichen ihre Ergebnisse auf sechs verschiedenen Bildersequenzen aus der Koloskopie.

1 Einleitung

Die Endoskopie ist als minimalinvasives Verfahren zur Diagnose und Therapie einer großen Anzahl an Krankheiten im klinischen Alltag von großer Bedeutung. Für die Diagnose von Dickdarmkrebs ist die Koloskopie (Darmspiegelung) das Standardverfahren, welches in Deutschland ab dem 55. Lebensjahr von den Krankenkassen alle zehn Jahre empfohlen und finanziert wird. Darmkrebs ist hierzulande bei beiden Geschlechtern die zweithäufigste Krebserkrankung und es werden im Jahr ca. 73.000 Neuerkrankungen registriert [1].

Während der Untersuchung werden digitale Momentaufnahmen zur Dokumentation und ggf. zur computergestützten Diagnose und Therapie [2] erstellt. Die Aussagekraft dieser Aufnahmen hängt nicht nur von der Qualität des bildgebenden Systems, sondern auch von Geschicklichkeit und Erfahrung des Untersuchers ab. Während einer Koloskopie bewegt sich der Darm durch die Peristaltik unentwegt, was die Aufnahme scharfer Einzelbilder erschwert. Ein Verfahren, welches vom Arzt gestartet, selbständig die schärfste Aufnahme aus einer frei wählbaren Anzahl von Folgebildern wählt, könnte hier verhindern, dass der Untersucher die Aufnahme mehrfach wiederholen muss und gleichzeitig die Qualität der Aufnahmen steigern. Aus diesen Gründen untersuchen wir verschiedene Schärfemaße im Echtzeiteinsatz auf Videos aus der Koloskopie.

Der Abschnitt 2 beschäftigt sich mit der Schärfe von Bildern und deren Messung. Hier werden verschiedene Schärfemaße und die durchgeführten Untersuchungen vorgestellt. In Abschnitt 3 werden die Resultate der Untersuchungen und die benötigten Rechenzeiten dargelegt. Im abschließenden Abschnitt 4 werden Schlussfolgerungen gezogen und ein Ausblick auf zukünftige Arbeiten gegeben.

2 Material und Methoden

Schärfe bezeichnet die Eigenschaft eines Bildes, Details an Kanten und Ecken genau darstellen zu können. Dies ist mit Hilfe lokal hoher Ortsfrequenzen möglich. Es gibt Verfahren, welche die Schärfe in Bildern bewerten. Die Farbwerte werden hierzu in Graustufen umgewandelt. Drei echtzeitfähige Verfahren werden hier vorgestellt und miteinander in der Anwendung auf Dickdarmpolypen unter NBI-Beleuchtung [3] verglichen, da diese Beleuchtung die diagnostisch wichtigen Blutgefäße [4] auf den Polypen hervorhebt. Im Folgenden werden die Schärfemaße Varianz, Sum Modulus Difference (SMD) und Perceptual Sharpness Metric (PSM) beschrieben.

2.1 Varianz

Im Jahr 1982 haben Erasmus et al. [5] vorgeschlagen, die über das gesamte Bild gemessene Varianz als Referenz für die Schärfe eines Bildes zu nutzen. Die Berechnung dieses Schärfemaßes erfolgt mit

$$S_{\text{Varianz}} = \frac{1}{N \cdot M} \sum_{i=1}^M \sum_{j=1}^N \left(x_{i,j} - \frac{1}{N \cdot M} \sum_{i=1}^M \sum_{j=1}^N x_{i,j} \right)^2 \quad (1)$$

wobei N der Bildbreite in Pixeln und M der Bildhöhe in Pixeln entspricht. Die Grauwerte der einzelnen Bildpunkte (i, j) werden mit $x_{i,j}$ bezeichnet.

2.2 Sum Modulus Difference

Sum Modulus Difference (SMD) [6] bewertet die partiellen Ableitungen nach x und y mit Hilfe der Amplitudendifferenzen von horizontalen und vertikalen Nachbarpixeln und summiert diese betragsmäßig auf. Das Schärfemaß wird anhand der Vorschrift

$$S_{\text{SMD}} = \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} (|x_{i,j} - x_{i-1,j}| + |x_{i,j} - x_{i,j-1}|) \quad (2)$$

berechnet.

2.3 Perceptual Sharpness Metric

Yang et al. [7] schlugen 2006 ein Schärfemaß namens Perceptual Sharpness Metric (PSM) vor. Das Bild wird überlappungsfrei in 8×8 -Pixel große Blöcke zerlegt. Auf diese Blöcke wird die Diskrete Kosinus-Transformation (DCT) wie in

$$DCT_{m,n}(k_1, k_2) = \sum_{i=0}^7 \sum_{j=0}^7 x_{i,j} \cos \left[\frac{\pi}{8} \left(i + \frac{1}{2} \right) k_1 \right] \cos \left[\frac{\pi}{8} \left(j + \frac{1}{2} \right) k_2 \right] \quad (3)$$

angewendet, wobei $k_1 = 0..7$ und $k_2 = 0..7$ die Indices für die Koeffizienten der DCT des transformierten 8×8 -Blockes mit der Position (m, n) darstellen. Die Koeffizienten der ersten Zeile und ersten Spalte des transformierten 8×8 -Blockes werden mit Hilfe der Sharpness Sensitivity Function $SSF(d) = d^{0.269}(-3.533 + 3.533d) \exp(-0.548d)$, welche auf Studien über die Empfindlichkeit des menschlichen Auges basiert [7], gewichtet und nach der Vorschrift

$$\tilde{S}_{m,n} = \sum_{d=1}^8 SSF(d) [DCT_{m,n}(0, d-1) + DCT_{m,n}(d-1, 0)] \quad (4)$$

aufsummiert. Abschließend wird die Summe

$$S_{PSM} = \frac{1}{M \cdot N} \sum_{m=1}^{M/8} \sum_{n=1}^{N/8} \tilde{S}_{m,n} \quad (5)$$

über alle Blöcke gebildet, wobei das Gesamtergebnis S_{PSM} auf die Bildbreite N in Pixeln und die Bildhöhe M in Pixeln normiert wurde.

2.4 Versuchsaufbau

Um die vorgestellten Verfahren zu vergleichen, haben wir sechs verschiedene Sequenzen aus endoskopischen Videos verwendet. Die Sequenzen zeigen verschiedene Dickdarmpolypen unter NBI-Beleuchtung. Die Videoausschnitte wurden zur einfacheren Handhabbarkeit in Halbbilder der Größe 1440×540 zerlegt. Die einzelnen Sequenzen sind zwischen 50 und 100 Bilder lang. Wie in der Koloskopie üblich zeigen viele Bilder Lichtreflexionen an der Darmoberfläche. Diese Glanzlichter wurden durch einen von Stehle im Jahr 2006 vorgestellten Algorithmus [8] lokalisiert und während der Berechnung der Schärfemaße ausgespart.

Verwendet wurde für die Testreihen ein Intel Core2 Duo E6550 mit 1 GB DDR2 RAM und NVIDIA GeForce 8400GS. Die Verfahren haben ihre Berechnungen in Echtzeit auf den zentralen Bildbereichen (576×468) ausgeführt, die Bilder wurden zur Dokumentation auf der Festplatte gespeichert und das beste Bild ausgewählt. Im Routineeinsatz reicht die Speicherung des derzeit besten Bildes im Speicher. Zu den drei Ergebnisbildern der Schärfemaße wurden scharfe Bilder aus den Serien herausgesucht und beigemischt. Diese Gruppe von Bildern für jede Serie wurde im Anschluss sieben in der Bildverarbeitung tätigen

Tabelle 1. Die Tabelle zeigt den prozentualen Anteil p der gewonnenen Entscheidungen für jedes Verfahren und jede Serie. In der vorletzten Spalte ist der Mittelwert jedes Verfahrens über alle Serien aufgeführt, in der letzten Spalte erscheint die Varianz.

	Serie 1	Serie 2	Serie 3	Serie 4	Serie 5	Serie 6	Mittelwert	Varianz
p_{Varianz}	0,64	0,43	0	0,14	0,45	0,49	0,36	0,057
p_{SMD}	0,40	0,76	0,60	0,66	0,68	0,12	0,54	0,056
p_{PSM}	0,52	0,03	0,45	0,19	0,04	0,32	0,26	0,043

Tabelle 2. Vergleich der durchschnittlichen Rechenzeiten T pro Bild in ms für die verschiedenen Serien.

	Serie 1	Serie 2	Serie 3	Serie 4	Serie 5	Serie 6	Mittelwert	Varianz
T_{Varianz}	3,77	4,06	3,92	3,65	3,78	3,92	3,85	0,02
T_{SMD}	12,07	12,04	12,50	13,60	12,78	11,73	12,45	0,45
T_{PSM}	19,21	19,84	20,28	21,37	19,68	20,05	20,07	0,54

Betrachtern vorgelegt. Die Aufgabe bestand darin, im direkten Vergleich der Bilder das jeweils schärfere zu bestimmen. Falls der Betrachter beide Bilder für gleich scharf hielt, wurde die Wertung geteilt. Die Gesamtsumme der von einem Verfahren erreichten Wertungen innerhalb einer Serie wurde dann auf die Anzahl der durchgeführten Vergleiche pro Bild normiert. Der Mittelwert aller Serien wurde dann genutzt, um das insgesamt beste Verfahren zu bestimmen.

3 Ergebnisse

Die drei Verfahren haben in den sechs Sequenzen jeweils das vermeintlich beste Bild ausgewählt. In der Abb. 1 sind beispielhaft die aus der Serie 6 gewählten Bilder für die drei Verfahren dargestellt. In Tabelle 1 sind die Wertungen aus der Studie zusammen gefasst. Die beste Gesamtbewertung durch die Betrachter erzielte das Verfahren SMD. Ebenfalls gemessen wurden die Rechenzeiten der einzelnen Verfahren, welche in Tabelle 2 dargestellt sind. Die niedrigste Rechenzeit T weist das Verfahren Varianz auf.

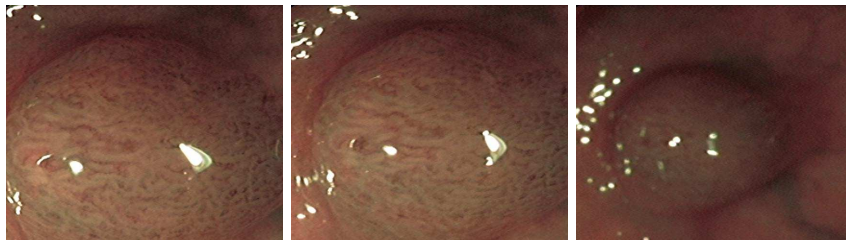


Abb. 1. Beispielhaft zeigt die Grafik die gefundenen Resultate von Varianz (Frame 6, links), SMD (Frame 17, Mitte) und PSM (Frame 60, rechts).

4 Diskussion

Das Ergebnis der Studie zeigt, dass die Bilder, die vom Verfahren SMD bestimmt wurden, von den Betrachtern als schärfer empfunden wurden als die Bilder der anderen Verfahren. Die Rechenzeiten T aller Verfahren zeigen deutlich, dass ihre Berechnung in Echtzeit möglich ist. Bis zu 50 Bilder pro Sekunde sind selbst mit PSM möglich. Allerdings schneidet dieses Verfahren erheblich schlechter ab als SMD, welches auch deutlich geringere Rechenzeiten ausweist. Wenn die Bilder als Grundlage für weiterführende Analysen [2] in Echtzeit genutzt werden, kann die Varianz Rechenzeit sparen. Im Falle der Gewinnung von Standaufnahmen senkt die Anwendung der Technik den Arbeitsaufwand für den untersuchenden Arzt, da weniger untaugliche Aufnahmen entstehen. Gleichzeitig steigert sich die Qualität der Aufnahmen für Dokumentation und Weiterverarbeitung.

In weiteren Untersuchungen sollte das Verfahren im Rahmen einer klinischen Studie getestet werden. Auf diesem Wege könnte die Qualität der gewonnenen Bilder direkt mit den Ergebnissen der herkömmlichen Standbilder verglichen werden. Weiterhin sollte untersucht werden, ob es möglich ist, zusätzliche Informationen über den Bildinhalt aus den 1-2 Sekunden Video zu extrahieren und in einem scharfen Standbild zu vereinen. Ein möglicher Ansatz wäre Extended Depth of Field (EDoF) [9].

Danksagung. Das Projekt wurde aus Mitteln der Exzellenzinitiative des Bundes und der Länder gefördert.

Literaturverzeichnis

1. Gesundheitsberichterstattung des Bundes. Krebs in Deutschland 2003-2004 - Häufigkeiten und Trends. 8th ed. Robert Koch Institut, Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V.; 2008.
2. Stehle T, Gross S, Behrens A, et al. Classification of colon polyps in NBI endoscopy using vascularization features. *Proc SPIE*. 2009;6918:12.
3. Gono K, Obi T, Yamaguchi M, et al. Appearance of enhanced tissue features in narrow-band endoscopic imaging. *J Biomed Opt*. 2004;9(3):568-77.
4. Tischendorf JJW, Wasmuth HE, Koch A, et al. Value of magnifying chromoendoscopy and narrow band imaging (NBI) in classifying colorectal polyps: A prospective controlled study. *Endoscopy*. 2007;39(12):1092-6.
5. Erasmus SJ, Smith KCA. An automatic focusing and astigmatism correction system for the SEM and CTEM. *J Microsc*. 1982;127:185-199.
6. Brenner JF, Dew BS, Horton JB, et al. An automated microscope for cytologic research. *J Histochem Cytochem*. 1976;24(1):100-11.
7. Yang KC, Guest CC, Das P. Perceptual sharpness metric (PSM) for compressed video. In: *Proc. IEEE ICME*; 2006. p. 777-80.
8. Stehle T. Removal of specular reflections in endoscopic images. *Acta Polytechn J Adv Eng*. 2006;46(4):32-6.
9. Bradley AP, Bamford PC. A one-pass extended depth of field algorithm based on the over-complete discrete wavelet transform. In: *Proc Image Vis Comput*; 2004. p. 279-84.

3D-Visualisierung und Kolokalisation von Proteinen und ceramidreichen Domänen

Christian Imhäuser^{1,2}, Heike Gulbins (Grassmé)², Erich Gulbins²,
Hans-Gerd Lipinski¹

¹Fachbereich Medizinische Informatik, Fachhochschule Dortmund

²Institut für Molekularbiologie, Universität Duisburg/Essen

ci@cne.de

Kurzfassung. Die Topologie von Proteinen innerhalb ceramidreicher Membranbereiche stimulierter T-Lymphozyten ist für die Molekularbiologie von großem Interesse, weil experimentell ein Zusammenhang zwischen der Clusterbildung von Rezeptorproteinen und der Membranaufteilung in ceramidreiche Mikrodomänen für den transmembranen Signalaustausch nachgewiesen werden konnte. Die Mechanismen der Clusterbildung und die räumliche Verteilung der Proteine sind bis heute nahezu unbekannt. Mit der modernen Laserscanmikroskopie lassen sich jedoch räumliche Bilddaten von Zellproben gewinnen, deren Bestandteile durch verschieden fluoreszierende Farbstoffe visualisiert werden. Mit Hilfe von adaptierten 3D-Visualisierungsmethoden konnten Volumerending- und Oberflächenmodelle erzeugt werden, anhand derer die räumliche Verteilung der Proteine und der ceramidreichen Bereiche sowie die Kolokalisation beider Moleküle dargestellt werden konnten.

1 Einleitung

Die Funktionen von Proteinen in lebenden Zellen werden unter Anderem durch ihre räumliche Verteilung festgelegt. Die räumliche Anordnung wird wiederum durch die Aufteilung in Membrandomänen beeinflusst. Bislang konnte gezeigt werden, dass durch die Aktivierung der sauren Sphingomyelinase es zu einer Formation ceramidreicher Membranbereiche kommt, die zu einer Clusterbildung aktivierter Rezeptormolekülen wie CD95 oder CD40 führt [1]. Funktioniert diese Clusterbildung aus verschiedenen Gründen nicht mehr, kann die Zelle weder zur Apoptose veranlasst, noch durch ein pathogenes Bakterium infiziert werden [2, 3]. Die entartete Zelle würde im ersten Fall ungehemmt wachsen und könnte im Gesamtorganismus zu einer krebsartigen Erkrankung führen.

Bis heute sind die exakten Mechanismen der Clusterbildung und die räumliche Verteilung der Proteine innerhalb der ceramidreichen Domänen nahezu unbekannt. Es werden daher geeignete 3D-Visualisierungsmethoden vorgestellt, die helfen, den exakten Prozess der Clusterbildung von Rezeptormolekülen sowie die räumliche Assoziation (Kolokalisation) mit den ceramidreichen Bereichen besser zu verstehen und interaktiv zu analysieren.

2 Material und Methoden

2.1 Bildakquise

Das verwendete Bildmaterial stammte aus einer schichtweisen Digitalisierung mehrerer Jurkat-Cluster (Zelllinien aus einer T-Lymphozyten Leukämie), welche mit substanzspezifischen Antikörpern präpariert wurden. Für das zellgebundene Ceramid wurde ein roter und für das CD95-Protein ein grüner Farbstoff verwendet. Die in den Antikörpern enthaltenden fluoreszierenden Farbstoffe wurden durch die Laserbestrahlung eines Konfokallasermikroskops optisch angeregt und ihre Intensität von einem Sensor erfasst. Die gewonnenen Bildinformationen wurden in Abhängigkeit von der roten bzw. grünen Farbe der Fluoreszenzstoffe in zwei separate Kanäle gespeichert.

Für die mikroskopischen Aufnahmen kam ein Laserscanmikroskop der Firma Leica DMIRE 2 (Öl benetztes Objektiv mit 100-facher Vergrößerung) zum Einsatz. Jede Zellprobe wurde mit einem Abstand zwischen $0,66$ und $0,91 \mu\text{m}$ abgetastet und wies durchschnittlich eine Größe von ungefähr $40 \times 40 \times 4 \mu\text{m}^3$ auf. Die akquirierten Bilder hatten eine Auflösung 512×512 Bildpunkten mit einer Pixelgröße von ca. $0,08 \mu\text{m}$.

2.2 Bildvorverarbeitung

Die einzelnen Bilder wiesen neben dem aufnahmebedingten typischen Rauschen Kontrastarmut und unausgewogene Helligkeitsverteilungen auf. Die geringe Tonwerttraumausnutzung im Original wurde zur Kontrastanhebung durch ein RescaleIntensityFilter korrigiert. Durch die Anzahl der Anhaftungen fluoreszierender Antikörper wurde die Intensität bei der Bildakquisition bestimmt. Ferner

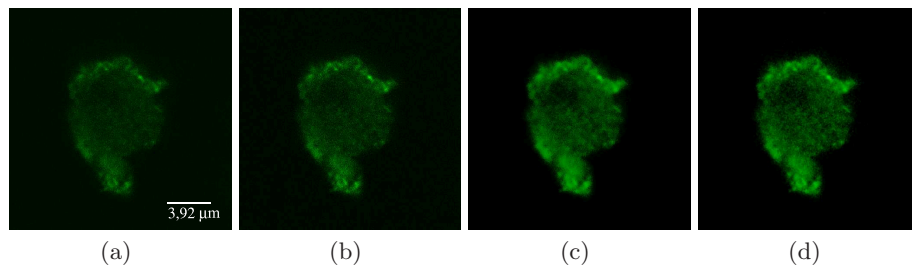


Abb. 1. Bildvorverarbeitung. (a) Exemplarisch gewähltes Schnittbild aus dem Jurkat-Cluster „CerCD952min1“. Ungefiltertes Einzelkanalbild mit verschiedenen Intensitäten von CD95 (grün). (b) Nach der Kontrastverbesserung. (c) Nach logarithmischer Skalierung und bilateraler Filterung (Bereich-Sigma = 30, Weite-Sigma = 40). Gute Rauschunterdrückung bei hervorragender Kanten-/Strukturerhaltung. (d) Nach logarithmischer Skalierung und Anwendung des Diffusionsfilters (Zeitschritt = 0.125, Leitwert = 6, Wiederholungen = 2). Gute Rauschunterdrückung bei hervorragender Kantenerhaltung. Innere Strukturen werden leicht geglättet.

ergab sich das Problem, dass sich Antikörper mit bereits angedockten verbinden konnten, d.h. je höher die Zahl der Anhaftungen, desto größer war auch die Wahrscheinlichkeit dieser vermeintlichen Intensitätsvergrößerungen. Daraus ließ sich kein linearer Zusammenhang zwischen akquirierter und realer Intensität mehr ableiten. Diskutiert wurde ein exponentieller Zusammenhang. Daher wurden die gewonnenen Bildinformationen logarithmisch skaliert, so dass der vermutete exponentielle Tonwertverlauf aufgehoben/umgekehrt wurde. Zusätzlich wurden Kontrast und Helligkeit unter Verwendung eines eigens entwickelten Filters erneut so angepasst, dass Werte außerhalb des definierten Tonwertraumes auf das jeweils Minimum/Maximum abgebildet wurden.

Um das Hintergrundrauschen weitgehend zu beseitigen, kam sowohl ein BilateralFilter als auch CurvatureAnisotropicDiffusionFilter zum Einsatz. Beiden Filtern ist es gemein, dass sie schwache Gradienten (wie Rauschen etc.) gut unterdrücken, aber große Steigungen der Intensitäten, welche oft an Kanten von Objekten innerhalb des Bildes vorzufinden sind, fast nicht beeinflussen.

2.3 Visualisierung

Die 3D-Visualisierungen der einzelnen Zellproben wurden mit Hilfe des Volumerenderings (RayCasting und TextureMapping) und der Oberflächenrekonstruktion nach dem MarchingCube-Verfahren erzeugt. Durch manuelle maßstabsgetreue Korrektur des Schnittbildabstandes wurden weitere Schnittbilder zwischen den gegebenen linear interpoliert. Farbtransfer- und Opazitätsfunktionen wurden so gewählt, dass die Intensitätsunterschiede mit denen im Bildmaterial übereinstimmten.

2.4 Kolokalisation

Für den visuellen Nachweis der Kolokalisationen von CD95 und Ceramid wurden mehrere verschiedene Methoden entwickelt oder angepasst. Die 3D-Daten wurden hierfür als gerastert in 3D-Zellen angesehen. Innerhalb dieser Zellen wurden alle Intensitätswerte oberhalb eines festgelegten Schwellwertes gezählt. Die minimalen und maximalen Z-Positionen wurden separat gespeichert. Bei der Produktintensität wurde die 3D-Zelle mit der Intensität aus dem normalisierten Produkt beider Summen vollständig, aber höchstens zwischen den vorher ermittelten Z-Positionen, gefüllt. Die Methode Binärisierung hingegen setzte die 3D-Zelle erst mit der maximalen Intensität, wenn beide Summen separat festgelegte Schwellwerte in relativem Maß überstiegen. Zudem wurde eine Kreuzkorrelationsfunktion (KKF) zur Quantifizierung des Kolokalisationsgrades von CD95 und Ceramid bestimmt. Da Korrelationen in Richtung der Z-Achse nicht von Interesse waren, wurde die zweidimensionale KKF schichtweise über den gesamten Bildstapel berechnet und als 3D-Objekt mit identischem Schnittbildabstand in die 3D-Szene eingebettet. Zusätzlich wurden die Werte der KKF mit Hilfe des RescaleIntensityFilters normalisiert.

Bis auf die eigens entwickelten Kolokalisationsroutinen wurden Algorithmen und Methoden des Insight Toolkits (ITK)[4] und Visualization Toolkits (VTK)[5]

unter der Entwicklungsumgebung Microsoft Visual Studio 2008 Express Edition verwendet.

3 Ergebnisse

Die verwendeten Schnittbildfolgen konnten im Hinblick auf Kontrastverbesserung, korrigierter logarithmischer Skalierung und Rauschunterdrückung für eine optimale 3D-Visualisierung aufbereitet werden. Die Abb. 1 zeigt die visuellen Ergebnisse nach der schrittweisen Filteranwendung auf ein exemplarisch gewähltes Schnittbild. Die räumliche Verteilung von CD95-Proteinen und Ceramid konnte unter Verwendung drei verschiedener Visualisierungstechniken zum ersten Mal gezeigt werden. Die Abb. 2(a-c) gibt neben der Oberflächenrekonstruktion auch die Visualisierungen mit Hilfe der beiden Volumerending-Verfahren wieder, anhand derer wir die verschiedenen Intensitätsverteilungen der Zellbestandteile

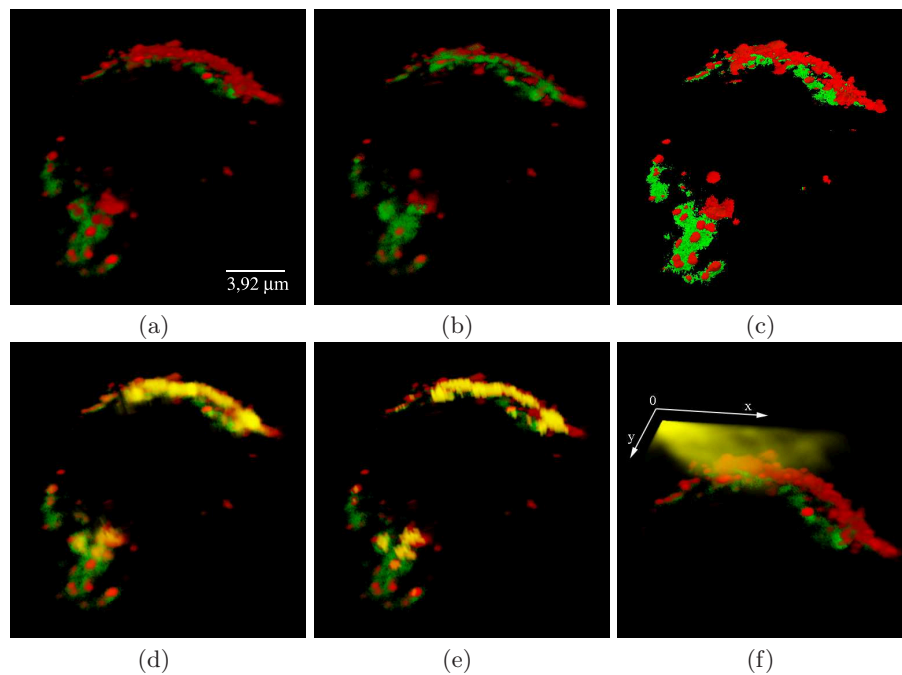


Abb. 2. 3D-Visualisierungen des Jurkat-Clusters „CerCD952min1“ als Overlay (CD95 = grün, Ceramid = rot, Kolokalisation = gelb). Räumliche Intensitätsverteilung mit Hilfe des (a) RayCasting- und (b) TextureMapping3D-Verfahrens (Schwellwerte für CD95/Ceramid = 128/103), und (c) Oberflächenrekonstruktion (Schwellwerte für CD95/Ceramid = 147/157). 3D-Darstellung der Kolokalisationen im Volumerending-Modell aus (a) unter Verwendung der Produktintensität (d) und Binärisierung (e). Dreidimensionale Kreuzkorrelationsfunktion beider Kanäle als eine Art „Ähnlichkeitswolke“ (f) angewandt auf den oberen Ausschnitt des Modells.

interpretieren konnten. Darüber hinaus wurden für den Nachweis der Kolokalisation die Verfahren Produktintensität und Binärisierung entwickelt und an acht Zellprobendaten getestet. Die Abb. 2(d-f) zeigt die visuellen Resultate der einzelnen Methoden, angewandt auf eine exemplarisch gewählte Zellprobe.

4 Diskussion

Die Bildanalyse ergab, dass mit den ausgewählten Filtern das Bildmaterial zugunsten einer optimalen 3D-Visualisierung aufbereitet werden konnte. Hierfür steuerte der `RescaleIntensityFilter` zu einer erheblichen Kontrastverbesserung bei. Kaum wahrzunehmende Intensitäten wurden sichtbar gemacht und vorhandene Konzentrationen hervorgehoben. Leider wurde auch das aufnahmebedingte Rauschen verstärkt, welches mit der nachfolgenden logarithmischen Skalierung noch unterstützt wurde. Trotzdem konnte der `BilateralFilter` diese Störungen gut beseitigen, und erhielt dabei hervorragend Kanten und innere Strukturen. Im Vergleich dazu glättete die optimierte Variante des Diffusionsfilters bei guter Rauschunterdrückung leicht die inneren Zellstrukturen und beeinflusste damit die granularitäre und wichtige Verteilung des Ceramids. Die Kanten wurden nahezu vollständig erhalten.

Im Hinblick auf die räumliche Verteilung von CD95 und Ceramid stellte sich das `RayCasting` als die beste Visualisierungstechnik heraus, da sich mit ihr die räumliche Intensitätsverteilung sehr präzise darstellen ließ. Die Oberflächenrekonstruktion bot hingegen die performantere Interaktionsmöglichkeit bei Verlust der Intensitätsdarstellung.

Es wurden erfolgreich Bildanalysemethoden entwickelt, mit denen der räumliche Nachweis der Kolokalisation gelang. Durch die Einführung der Produktintensität konnte der Grad der Kolokalisation bestimmt werden. Der Ungenauigkeit hinsichtlich der multiplikativen Eigenschaft des Verfahrens wurde mit einer gewichteten Skalierung entgegengewirkt. Mit Hilfe der Binärisierung konnten für jeden Kanal die Positionen der Kolokalisation präziser bestimmt werden. Ebenso konnte die räumliche Darstellung der Kreuzkorrelationsfunktion als ein Ähnlichkeitsmaß interpretiert werden. Allerdings ließ sich mit ihr keine Aussage über Intensität und Position der Kolokalisation treffen.

Literaturverzeichnis

1. Grassmé H, Jekle A, Riehle A, et al. CD95 Signaling via ceramide-rich membrane rafts. *J Biol Chem.* 2001;276(23):20589–96.
2. Krawczyk C, Penninger JM. Molecular controls of antigen receptor clustering and autoimmunity. *Trends Cell Biol.* 2001;11(5):212–20.
3. Grassmé H, Jendrossek V, Riehle A, et al. Host defense against *Pseudomonas aeruginosa* requires ceramid-rich membrane rafts. *Nat Med.* 2003;9(3):322–30.
4. Ibáñez L, Schroeder W, Ng L, et al. *The ITK Software Guide*. 2nd ed. United States of America: Kitware Inc; 2005.
5. Schroeder W, Martin K, Lorensen B. *The Visualization Toolkit*. 4th ed. United States of America: Kitware Inc; 2006.

Schnelle Zugangsplanung für die perkutane Punktion der Leber

Markus Engel¹, Alexander Seitel¹, Markus Fangerau¹, Boris A. Redeleff²,
Christof M. Sommer², Caroline Essert-Villard³, Claire Baegert³,
Hans-Peter Meinzer¹, Lena Maier-Hein¹

¹Abt. für Medizinische und Biologische Informatik, DKFZ Heidelberg

²Abt. für Diagnostische und Interventionelle Radiologie, Radiol. Klinik Heidelberg

³Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection, Illkirch

m.engel@dkfz-heidelberg.de

Kurzfassung. Minimal-invasive Verfahren zur Krebsdiagnose und -therapie erfordern häufig das Einführen eines nadelförmigen Instrumentes in das Zielorgan. Um durch eine gut gewählte Nadeltrajektorie eine komplikationsfreie, schnelle Durchführung der Intervention zu erreichen, wurde in verwandten Arbeiten ein sogenanntes Constraint-Konzept für die automatische Zugangsplanung entwickelt, durch welches mögliche Insertionszonen auf der Haut berechnet und entsprechend einer Kostenfunktion bewertet werden. Um die Planung zu beschleunigen, stellen wir in diesem Beitrag eine grafikartenbasierte Implementierung des sogenannten Occlusion Constraints vor, welcher die Hautregionen als Insertionszonen ausschließt, deren Wahl mit der Durchquerung einer Risikostruktur verbunden wäre. Desweiteren zeigen wir retrospektiv anhand von sechs klinischen Fällen, bei denen eine suboptimale Nadeltrajektorie zu z.T. schweren Komplikationen führte, dass die Sicherheit für den Patienten durch eine automatische Planung drastisch erhöht werden kann.

1 Einleitung

Minimal-invasive Verfahren zur Krebsdiagnose und -therapie (z.B. Nadelbiopsien, Radiofrequenzablationen) lösen offene Verfahren aufgrund ihres schonenden Charakters in der klinischen Routine zunehmend ab. Diese Eingriffe erfordern typischerweise das Einführen eines nadelförmigen Instrumentes in das Zielorgan. Eine gut gewählte Nadeltrajektorie ist dabei essentiell für eine komplikationsfreie, schnelle Durchführung der Intervention. Die in der Literatur vorgestellten Ansätze zur computergestützten Planung beschränken sich hauptsächlich auf semi-automatische Methoden [1, 2, 3], welche eine gute Einschätzung gegebener Pfade seitens des Arztes voraussetzen. Die unserer Kenntnis nach einzige vollautomatische Methode wurde von Baegert et al. [4] publiziert. Die Autoren ermitteln mögliche Insertionszonen auf der Haut mittels sogenannter Hard Constraints und bewerten deren Güte anhand sogenannter Soft Constraints (Abs. 2). Um die Planung zu beschleunigen, stellen wir in diesem Beitrag eine grafikartenbasierte Implementierung des Occlusion Constraints vor, welcher diejenigen

Hautregionen als Insertionszonen ausschließt, deren Wahl mit der Durchquerung einer Risikostruktur verbunden wäre. Desweiteren führen wir einen zusätzlichen Soft Constraint ein, der Trajektorien bevorzugt, die in der axialen CT-Schicht liegen. Schließlich zeigen wir retrospektiv anhand von sechs klinischen Fällen, bei denen eine suboptimale Nadeltrajektorie zu z.T. schweren Komplikationen führte, dass die Sicherheit für den Patienten durch eine automatische Planung drastisch erhöht werden kann.

2 Methoden

Das Constraint-Konzept zur automatischen Pfadplanung wird detailliert in [4] beschrieben und ist in Abb. 1 zusammengefasst. Um die Tatsache zu berücksichtigen, dass Trajektorien, die nahezu in einer axialen Schicht liegen, aufgrund einer leichteren Umsetzung der Planung vom Arzt bevorzugt werden, wurde der sog. In-Plane Constraint zu den Soft Constraints hinzugefügt. Desweiteren wurde eine effiziente Implementierung des Occlusion Constraints zwecks Laufzeitoptimierung entwickelt, wie im folgenden Abschnitt beschrieben wird.

2.1 Laufzeitoptimierung durch asynchrones Ansprechen der GPU

Die von Baegert et al. verwendete (nicht publizierte) Implementierung des Occlusion Constraints (OC) basiert auf den sog. Occlusion Queries (OQ). Dabei handelt es sich um ein Feature neuer Grafikkarten, das es ermöglicht, die Anzahl gezeichneter Pixel einer Struktur beim Rendering der Szene zu zählen. Um dies für die Pfadplanung zu nutzen, wird die (virtuelle) Kamera der Szene in den geplanten Zielpunkt (i.d.R. das Tumorzentrum) gesetzt und anschließend jedes Oberflächendreieck zwei mal gezeichnet (Abb. 2). Im ersten Durchgang (*Z-Puffer Überprüfung aus*) erhält man für das aktuelle Polygon die maximal zeichenbare Anzahl an Pixeln (entspricht keiner Verdeckung). Im zweiten Rendering (*Z-Puffer Überprüfung an*) erhält man die Anzahl nicht verdeckter, also von der Kameraposition aus sichtbarer Pixel. Die Differenz der beiden Werte dient zur Berechnung der Verdeckung des Polygons wie in Abb. 2 illustriert. Obwohl dieser Ansatz schon von der hohen Leistungsfähigkeit moderner Grafikkarte profitiert, besteht der Nachteil, dass CPU und GPU nicht parallel arbeiten: Die CPU wartet auf die GPU, während diese ein Polygon zeichnet, und die GPU wartet auf die CPU, während diese die Ergebnisse der OQ auswertet.

Zwecks Laufzeitoptimierung haben wir das Verfahren wie folgt erweitert: Initial wird ein Puffer von OQ angelegt, der solange ergänzt wird, bis alle Dreiecke gerendert wurden. Parallel liest die CPU nach und nach die Ergebnisse der bereits beendeten OQ aus, verarbeitet diese und leert sukzessive den Zwischenspeicher. Beide Einheiten können ihre Ressourcen dadurch maximal ausnutzen.

2.2 Evaluation

Die Evaluation der automatischen Pfadplanung wurde retrospektiv anhand von sechs klinischen Fällen (fünf Radiofrequenzablationen, eine Abszesspunktion)

durchgeführt, bei denen eine suboptimale Nadeltrajektorie zu schweren Komplikationen (Pneumothorax) führte. In jedem Datensatz wurden Tumor, Leber und Risikostrukturen (Abb. 1) segmentiert, und der geplante Zielpunkt wurde

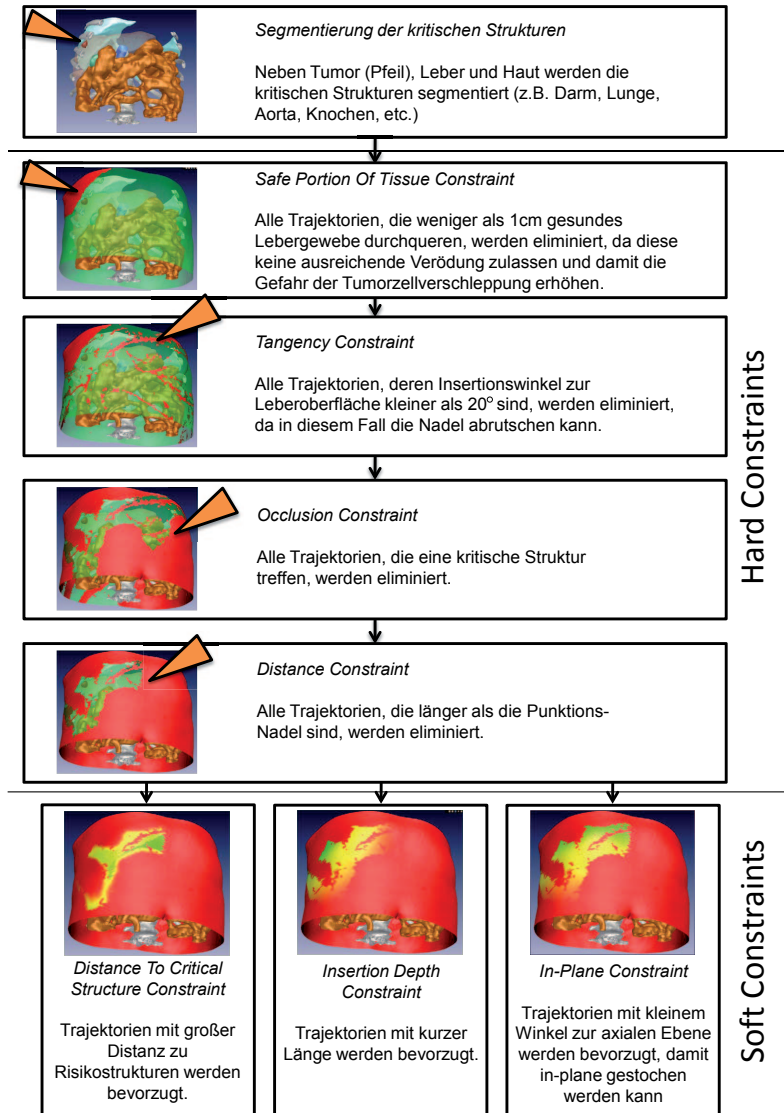


Abb. 1. Schematischer Ablauf der automatischen Pfadplanung. Nach der Segmentierung der Daten eliminieren die Hard Constraints alle verbotenen Bereiche der Haut (rot). Anschließend werden die verbleibenden Zonen durch die Soft Constraints bewertet und farblich kodiert (rot→grün).

Tabelle 1. Vergleich der durchschnittlichen Laufzeiten der einzelnen Constraints und deren Auswirkung auf die Fläche der verbleibenden, erlaubten Insertionszonen.

Constraint	ØLaufzeit [ms]	erlaubte Fläche isoliert [%]	erlaubte Fläche in Pipeline [%]
Safe Portion Of Tissue	1308 ±144	98 ±2	98 ±2
Tangency	93 ±12	98 ±0	96 ±5
Occlusion	2525 ±336	25 ±9	24 ±9
Distance	23 ±0	46 ±22	16 ±8

auf den Schwerpunkt des Tumors gesetzt. Zur Kompensation der Atembewegung wurde die Lunge um 1,2 cm in superior-inferior- und um jeweils 0,4 cm in anterior-posterior- und dexter-sinister-Richtung dilatiert. Die Performanz der anschließenden Pfadplanung wurde anhand der Laufzeit jedes Filters ermittelt. Desweiteren wurde der für den Eingriff tatsächlich gewählte Einstichspunkt auf den Planungsdatensatz registriert und ermittelt, ob dieser nach Berechnung der Hard Constraints zulässig gewesen wäre.

3 Ergebnisse

Tabelle 1 zeigt die durchschnittliche Laufzeit für jeden Hard Constraint, sowie den prozentualen Anteil der Hautoberfläche, der nach dessen Anwendung noch erlaubt ist. Die Laufzeit aller Hard Constraints liegt im Mittel unter 3 s.

Im Vergleich zu der von Baegert et al. verwendeten [4], synchronen Implementierung konnte eine Reduktion der Laufzeit um den Faktor 14,4 erreicht werden ($2,5 \pm 0,3$ s vs $36 \pm 8,9$ s). Die mittlere Laufzeit der Soft Constraints lag bei $9,1 \pm 5,4$ s (*Distance To Critical Structures*), $0,02 \pm 0$ s (*Insertion Depth und In-Plane*).

Laut der Evaluationsstudie wären drei der sechs in der Klinik gewählten Einstichspunkte unzulässig gewesen. Bei zwei weiteren Patienten war die Trajekto-

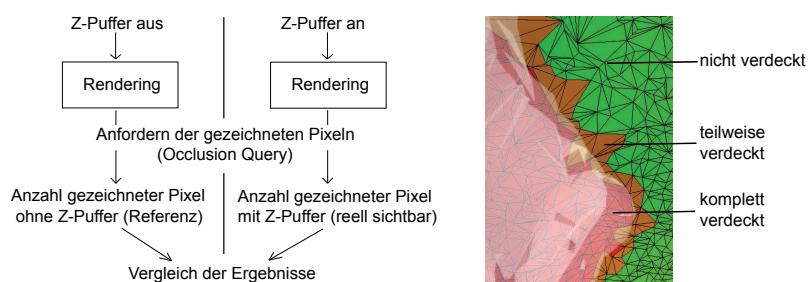


Abb. 2. Links: Schematische Darstellung der Implementierung der OC; rechts: Mesh der Haut farblich kodiert durch den OC mit transparenter Risikostruktur: komplett oder teilweise verdeckt → verboten; nicht verdeckt → erlaubt.

rie zwar zulässig, jedoch hätte es in beiden Fällen Insertionspunkte gegeben, die bzgl. *aller* Parameter besser bewertet worden wären. Die Distanz zu kritischen Strukturen konnte bis zu 0,3 cm (im Mittel 0,25 cm) bzw. 5,9 cm (im Mittel: 2,8 cm) erhöht und die Trajektorie um 0,03 cm (im Mittel 0,02 cm) bzw. 2,2 cm (im Mittel 0,9 cm) verkürzt werden, ohne den Winkel der Nadel zur axialen Ebene zu verschlechtern. Beim sechsten Datensatz konnten zwar nicht alle Parameter verbessert werden, jedoch hätte man mit einer Verlängerung der Trajektorie um nur 1,3 cm die Distanz zu Risikostrukturen um 12,7 cm vergrößern können, um so einen Pneumothorax auszuschließen.

4 Diskussion

Der Beitrag dieser Arbeit besteht in (1) der Erweiterung der in [4] bereits publizierten Constraints um den In-Plane Constraint, (2) einer nachweislich schnelleren Grafikkartenimplementierung des OC und (3) einer retrospektiven Evaluationsstudie auf klinischen Daten zur Demonstration der Notwendigkeit einer computergestützten Pfadplanung. Mit einer Gesamtlaufzeit von $13,1 \pm 5,3$ Sekunden ($n=6$) bietet die implementierte automatische Trajektorienplanung eine genügend geringe Laufzeit für den klinischen Einsatz. Allerdings sollte, wie in [4] angesprochen, eine schnelle, semi-automatische Methode zur Organsegmentierung benutzt werden, damit die Zeit für die Aufbereitung der Daten auf wenige Minuten reduziert werden kann. Bei Verwendung einer computergestützten Planung in der Klinik sollte unbedingt darauf geachtet werden, dass die präinterventionellen Daten im gleichen Zustand im Atemzyklus akquiriert werden, in dem auch die Nadelinserterion durchgeführt wird (typischerweise volle Expiration). Da diese Vorbedingung bei den im Rahmen dieser Studie untersuchten Daten nicht immer gegeben war, kam es vor, dass Einstichpunkte, die zu einem Pneumothorax führten, nicht als "unzulässig" durch den OC markiert wurden. Dennoch erhielten diese Punkte eine schlechte Bewertung aufgrund des Distance To Critical Structure Constraints, was für die Robustheit der Methode gegenüber atembedingter Organverschiebungen spricht. Zusammenfassend lässt sich feststellen, dass diese Studie den Einsatz einer Constraint-basierten Zugangsplanung in der Klinik nahelegt.

Literaturverzeichnis

1. Hirooka M, et al. Virtual puncture line in radiofrequency ablation for hepatocellular carcinoma of the caudate lobe. *Am J Radiol.* 2009;193(2):W149–W151.
2. Zhai W, et al. Preoperative surgery planning for percutaneous hepatic microwave ablation. *Lect Notes Computer Sci.* 2008;5242:569–77.
3. Butz T, et al. Pre- and intra-operative planning and simulation of percutaneous tumor ablation. *Lect Notes Computer Sci.* 2000;1935:11–4.
4. Baegert C, et al. Multi-criteria trajectory planning for hepatic radiofrequency ablation. *Lect Notes Computer Sci.* 2007;4792:676–84.

Non-Rigid Registration to Capture Optic Nerve Head Variability

Jan Paulus^{1,2}, Rüdiger Bock^{1,2}, Volker Daum¹, Joachim Hornegger^{1,2}

¹Pattern Recognition Lab, Martensstr. 3, 91058 Erlangen, Germany

²Graduate School in Advanced Optical Technologies (SAOT)

Friedrich-Alexander-University Erlangen-Nuremberg

`ruediger.bock@informatik.uni-erlangen.de`

Abstract. Glaucoma causes variations of the optic nerve head (ONH). We propose an approach to describe the inter subject variability of the ONH by dense deformation fields produced by non-rigid registration. Non-glaucoma but acquisition related variations are eliminated in a pre-processing step. Standard non-rigid registration was extended by a radial smoothing regularizer to preserve the ONH natural circular appearance and to enforce a physiological suitable mapping. The proposed approach captures successfully ONH variations.

1 Introduction

Glaucoma is one of the leading causes for blindness as it induces a degeneration of retinal nerve fibers. The occurrence of the disease is also accompanied by changes of the optic nerve head's (ONH) appearance. Based on a quantitative characterisation of the ONH several glaucoma indices were developed [1]. However, most of these techniques rely on geometric measurements or parameters of the ONH such as diameter or area. From our point of view, the established techniques only utilize a sparse characterization of the ONH.

Thus, we propose to describe the inter subject variability of the ONH by dense deformation fields produced by non-rigid registration. It represents the first step towards high resolved statistically modelling of ONH variability.

To the authors' knowledge existing registration approaches on retinal images are mostly used for mapping intra subject images in longitudinal studies or mosaicing-like methods. There are several feature based approaches utilizing landmarks out of the segmented vessel tree to map template and reference image. The mapping transformation is represented by a stand-alone quadratic model [2] or by an affine model that can be refined by a quadratic model if necessary [3, 4]. Region based approaches are used to e.g. reconstruct the ONH depth information from stereo images [5]. Different metrics and optimization methods for rigid registration are compared by Karali et al. [6].

2 Material and Methods

The proposed registration approach is divided into two steps. (i) Acquisition dependent translations, not related to the disease are excluded by a previous rigid registration. (ii) The variability of the ONH is captured by non-rigid registration that additionally accounts for a natural circular mapping and the individual vessel structure.

In the following the used state-of-the-art metrics and registration techniques are based on Modersitzki [7].

2.1 Elimination of Translation variations

The rigid registration is applied on gradient magnitude fundus images. This representation emphasizes the ONH as region of interest and suppresses the individual vessel tree in case of large filter support and is less sensitive to luminosity variations (Fig. 1).

2.2 Capturing the Papilla Shape Variability

In order to capture the inherent ONH variability, we apply an adapted non-rigid registration approach. It is extended by a regularisation term approving radial deformations to ensure a physiological mapping of the corresponding ONH sectors. Since the vessel tree is highly individual like a fingerprint it is masked out during the calculation of the distance measures. The registration problem is formulated as an constraint optimization problem

$$\min_{\mathbf{U}} J = D(\mathbf{R}, \mathbf{T}_{\mathbf{U}}) + \tau S(\mathbf{U}) \quad (1)$$

with the deformation field \mathbf{U} minimizing the distance D between the reference image $\mathbf{R} \in R^{n \times m}$ and the deformed template image $\mathbf{T}_{\mathbf{U}} \in R^{n \times m}$ while constrained by a smoothing factor S .

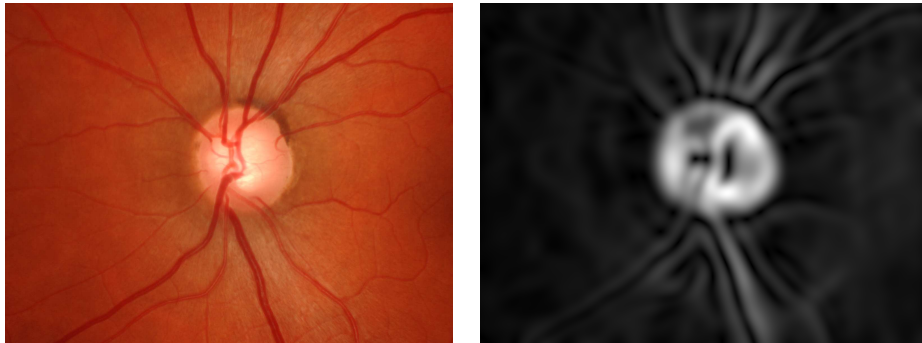


Fig. 1. Example for the preprocessing before applying rigid registration. The gradient magnitude image (right) computed from the input image (left) emphasizes the ONH as region of interest.

We additionally developed a so called radial smoothing S_{RS} to regularize the registration to favor a circular growing or shrinking of the ONH. The introduced weighted radial term measures the deviation of the deformation's directions to the radial directions originated in the ONH center (c_0, c_1)

$$S_{RS}(\mathbf{U}) = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left(\|\mathbf{u}(i, j)\|_{\epsilon}^2 - \left(\frac{\mathbf{u}^T(i, j)\mathbf{r}(i, j)}{\|\mathbf{r}\|_{\epsilon}} \right)^2 \right), \mathbf{r}(i, j) = \begin{pmatrix} i - c_0 \\ j - c_1 \end{pmatrix} \quad (2)$$

with $\|\cdot\|_{\epsilon}$ the modified Euclidean norm to prevent the result of being zero. S_{RS} preserves the circular shape of the ONH and prevents unnatural results like surface-water-like appearance.

3 Results

The registration was evaluated on 60 randomly selected image pairs with a size of 1600×1216 for each image. As gold standard, we used the translation between the ONH centers and the segmented ONH rims after registration automatically determined by an ONH segmentation technique [8].

3.1 Rigid Registration

As distance measures for the rigid registration we applied ‘‘Sum of Squared Distances (SSD)’’ and ‘‘Normalized Correlation (NC)’’ optimized by ‘‘Regular Step Gradient Descent Optimizer’’ to build a translation model. Both metrics are fast and sufficient for mapping two blob-like shapes of nearly similar intensities.

Initially, only 53.33% of the pairs showed an ONH distance lower than 150 pixels. This rate could be improved up to 76.67% for SSD and 78.33% for NC. Overall, the rigid registration on gradient magnitude filtered fundus images showed a reasonable compensation of translation variations.

3.2 Non-rigid Registration

For the the non-rigid registration, we evaluated ‘‘Sum of Squared Distances (SSD)’’ as mono-modal and ‘‘Mutual Information (MI)’’ as multi-modal metric to incorporate the impact of the luminosity inhomogeneities and the non-standardized intensities.

Fig. 2 shows an example of a deformed template image using SSD. It illustrates that the additional radial regularization is able to preserve the natural circular shape of the ONH in contrast to the bumped ONH in case of stand-alone SSD.

For a quantitative evaluation, we compared the automatically determined radii of the ONH rims of the reference and the deformed template images. This criteria is independent from the optimized distance measure. Tab. 1 shows the non-rigid registration utilizing SSD as distance measure results in the lowest radii difference average of around 34 pixels. As the SSD in combination with radial smoothing additionally decreases the standard deviation from 38.0 for stand-alone SSD to 33.1. We consider this configuration as most reliable.

Table 1. Average error and standard deviations for radius differences between reference and deformed template ONH in 60 randomly chosen image pairs. In contrast to mutual information (MI), the sum of squared differences (SSD) in combination with the proposed radial smoothing (RS) produces the lowest average error and standard deviation to be considered as a reliable technique for ONH variability capture.

Metric	Average error	Standard deviation
Initial	45.08	39.46
MI	41.95	38.97
MI + RS	43.60	40.69
SSD	34.52	38.00
SSD + RS	34.52	33.14

4 Discussion

The rigid registration is able to reduce the inter ONH distances. But the quality of the input images, mainly luminosity homogeneity, is still important. The

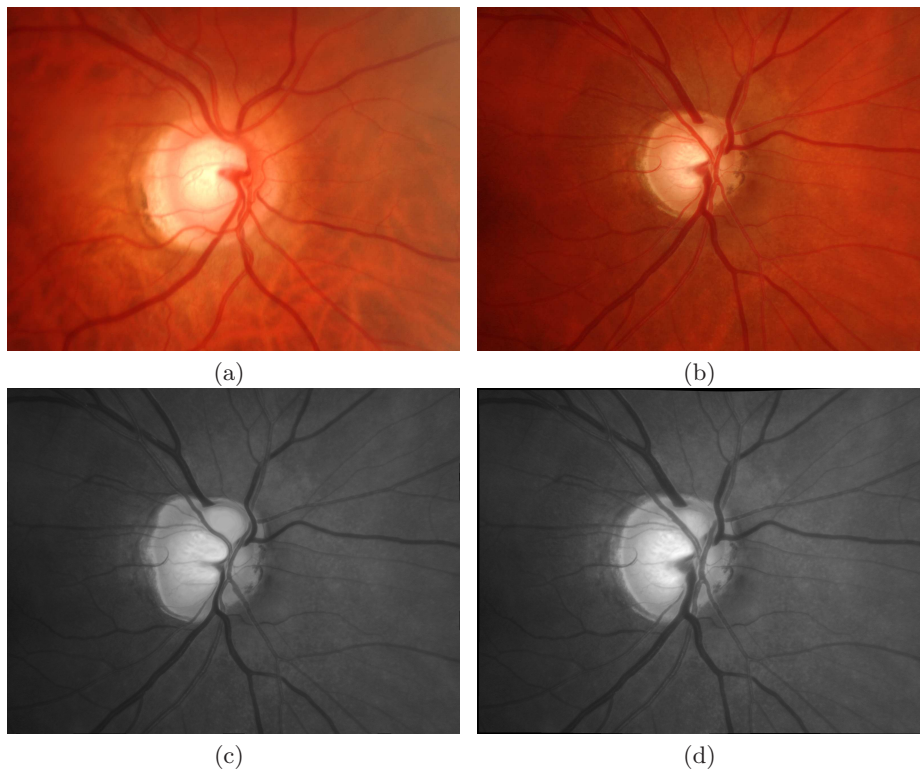


Fig. 2. Reference (a), template (b) and result of non-rigid registration using sum of square differences stand-alone (c) and in combination with radial smoothing (d) showing that the radial smoothing is able to preserve the natural circular optic nerve head appearance.

gradient based preprocessing emphasizes the ONH only at a border area, if the luminosity variations are too strong around the papilla, which results in less reduced distances after the registration. Nevertheless, translation variations can be reduced sufficiently for the ONH variability capturing.

The results of the non-rigid registration have to be considered as a trade-off between preserving the natural appearance of the template ONH and adapting the shape of the reference and the template ONH. After a visual inspection of the results SSD and MI stand-alone tend to produce results like surface-water-like appearances inside the ONH area or unnatural formed ONH borders.

Overall the proposed registration pipeline is able to reliably capture inter subject ONH variability as it clearly reduces the average difference between the ONH diameters of the reference and the deformed template images. Only due to the radial smoothing, related ONH sectors are physiological correctly matched and the deformed ONH retains its natural appearance.

This technique is able to densely describe the ONH variability by deformation fields and can be considered as the first step towards a highly resolved statistical deformation modeling to gain new insights to e.g. the glaucoma disease.

References

1. Sharma P, Sample PA, Zangwill LM, et al. Diagnostic tools for glaucoma detection and management. *Surv Ophthalmol.* 2008;53 Suppl1:S17–S32.
2. Baumgarten D, Doering A. Registrierung von aufnahmen des augenhintergrundes zur erstellung großflächiger kompositionsaufnahmen. *Biomed Eng.* 2007;52(6):365–74.
3. Can A, Stewart CV, Roysam B, et al. A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. *IEEE Trans Pattern Anal Mach Intell.* 2002;24(3):347–64.
4. Chanwimaluang T, Fan G, Fransen SR. Hybrid retinal image registration. *IEEE Trans Inf Technol Biomed.* 2006;10(1):129–42.
5. Ritter N, Owens R, Cooper J, et al. Registration of stereo and temporal images of the retina. *IEEE Trans Med Imaging.* 1999;18(5):404–18.
6. Karali E, Asvestas P, Nikita KS, et al. Comparison of different global and local automatic registration schemes: an application to retinal images. In: *Proc MICCAI.* vol. 1; 2004. p. 813–20.
7. Modersitzki J. *Numerical Methods for Image Registration.* Oxford: Oxford University Press; 2004.
8. Chrástek R, Wolf M, Donath K, et al. Automated segmentation of the optic nerve head for diagnosis of glaucoma. *Med Image Anal.* 2005;9(4):297–314.

Effiziente Methode zur Generierung von Ganzkörperdaten für die Fettgewebsanalyse

Diana Wald¹, Tobias Schwarz¹, Markus Fangerau¹, Julien Dinkel², Stefan Delorme², Rudolf Kaaks³, Hans-Peter Meinzer¹

¹ Medizinische und Biologische Informatik, DKFZ Heidelberg

² Radiologie, DKFZ Heidelberg

³ Epidemiologie von Krebserkrankungen, DKFZ Heidelberg

d.wald@dkfz-heidelberg.de

Kurzfassung. Ganzkörper-MRT findet im medizinischen Alltag immer mehr Anwendung in der Früherkennung und Ausbreitungsdiagnostik von Erkrankungen. Aufgrund eines erhöhten Rechenaufwandes bei bestimmten Aufnahmetechniken, wie beispielsweise dem Dixon Protokoll, kann die automatische Generierung der Ganzkörperdaten durch die Gerätekonsole nicht erfolgen, es werden nur die einzelnen Segmente gespeichert. In diesem Beitrag wird ein neues Verfahren zur Generierung artefaktfreier Ganzkörperdaten vorgestellt. Spezielles Augenmerk wird dabei auf das Zusammenfügen mehrerer Bildabschnitte gelegt, wobei die Übergänge geglättet werden, um die Qualität der Randbereiche zwischen den einzelnen Segmenten aufzubessern. Die Optimierung erfolgt durch eine Intensitätsnormalisierung und Interpolation zwischen den zusammengeführten Schichten. Weiterhin werden Artefakte der Dixon Sequenz automatisch ausgeglichen, indem Phasenfehler im Bild mittels Differenzbilder automatisch korrigiert werden.

1 Einleitung

Adipositas ist ein wichtiger Faktor für die Entstehung von chronischen Erkrankungen wie koronare Herzkrankheiten, Diabetes mellitus oder Krebs. Doll und Peto führten 1981 etwa 35% der Todesfälle an Krebs auf eine ungünstige Ernährung zurück [1]. Neue wissenschaftliche Studien haben gezeigt, dass ein Zusammenhang zwischen der Entstehung von chronischen Erkrankungen und dem Volumen des viszeralen Fettgewebes besteht [2]. Um weitere Studien mit Hilfe der medizinischen Bildverarbeitung zu unterstützen, werden für diesen Zweck hochaufgelöste Ganzkörper-Magnetresonanztomographen eingesetzt. Ganzkörper-MRT findet im klinischen Alltag nicht nur in der Analyse von Fettgewebe Anwendung, sondern bietet darüber hinaus eine medizinische Vorsorge zur Früherkennung und Ausbreitungsdiagnostik von Erkrankungen wie Entzündungen oder Tumoren [3].

Die Dixon Sequenz ist ein spezielles MRT Protokoll für die Fettgewebsanalyse und separiert das Fett- vom Wassergewebe, indem es die unterschiedlichen

Resonanzfrequenzen der Protonen in Fett und Wasser ausnutzt (Chemische Verschiebung) [4]. Durch einen erhöhten Rechenaufwand können die blockweise aufgenommenen Daten nicht auf der Konsole zusammen gesetzt werden und ein Nachbearbeitungsprozess wird notwendig. Das Hauptproblem bei der Erstellung eines Ganzkörperdatensatzes aus den einzelnen Segmenten bilden die stark veräuschten und intensitätsschwachen Bildränder. Bis heute sind wenige wissenschaftliche Arbeiten publiziert wurden, die sich mit der automatischen Generierung von Ganzkörperdaten befassen [5]. Zudem gibt es ein weiteres Artefakt der Dixon Sequenz: Bei diesem Protokoll kommt es häufig zu einem Phasenfehler im Bild, verursacht durch das inhomogene Magnetfeld. In der Regel greift hier der Unwrapping Algorithmus ein, der auf Region Growing basiert [4]. Jedoch schlägt diese Korrektur in den Beinen häufig fehl und wir erhalten eine fehlerhafte Ausgabe (Abb. 1(d,e)). Bisher gibt es noch keine wissenschaftliche Arbeit, die diesen markanten Fehler in einem Nachbearbeitungsprozess korrigiert.

Glocker und Wachinger haben in [5] eine Methode vorgeschlagen, die durch deformierbares Zusammenfügen der einzelnen Bildabschnitte die Bildrandartefakte korrigiert. Dabei werden überlappende Regionen der einzelnen Abschnitte verwendet. Anhand dieser Methode konnte eine Kompensation der Artefakte und eine Generierung hoch aufgelöster Ganzkörperdaten erzielt werden. Dennoch wird für das Verfahren eine Rechenzeit von etwa 25 min für drei Bildabschnitte benötigt, was auf den Bearbeitungsschritt der simultanen Registrierung zurückzuführen ist. Ziel dieser Arbeit ist die Entwicklung eines effizienten Verfahrens zur Generierung artefaktfreier Ganzkörperdaten. Unsere Methode verwendet Bildabschnitte, die sich nicht überlappen, um eine rechenaufwendige Registrierung zu vermeiden. Zusätzlich verbessern wir die Bildqualität an den Randbereichen durch eine Intensitätsnormalisierung und Interpolation zwischen den einzelnen Bildschichten. Das Verfahren generiert einen Ganzkörperdatensatz aus sechs bis sieben Bildabschnitten innerhalb von $\approx 10,13$ s und liefert einen glatten Übergang zwischen den einzelnen Segmenten. Neben der Generierung von Ganzkörperdaten stellen wir einen Prozess vor, der die Artefakte der Dixon-Sequenz entfernt: Falsch zugeordnete Regionen in den Fett- und Wasserbildern werden automatisch erkannt und korrigiert. Ergebnis dieser Arbeit sind artefaktfreie Ganzkörper-Fett-Daten, die innerhalb kürzester Zeit generiert werden und weitere Segmentierungen und Diagnosen zur Fettgewebsanalyse ermöglichen.

2 Material und Methoden

Die vorgestellten Verfahren werden anhand von Datensätzen von vier Probanden evaluiert. Die Dixon-Sequenz generiert automatisch in einem Aufnahmeprozess vier Ganzkörperdaten pro Proband (in-phase-, opposed-phase-, Fett- und Wasserdatensatz), somit kann eine Bewertung der Generierung der Ganzkörperdaten auf 16 Datensätzen erfolgen. Die Auflösung der Bildsegmente ist $384 \times 288 \times 64$ Voxel, mit einem Spacing von $1,3 \times 1,3 \times 3$ mm. Ein Ganzkörperdatensatz besteht aus sechs bis sieben Segmenten. Die Aufnahme erfolgte mit einem MRT

der Firma Siemens. Die Implementierung erfolgte innerhalb des Medical Imaging and Interaction Framework (MITK) [6].

2.1 Generierung von Ganzkörperdaten

Der Ansatz zur Generierung der Ganzkörperdaten verwendet nicht-überlappende Bildabschnitte. Die Anzahl der zu kombinierenden Segmente ist beliebig. Aufgrund der Ganzkörperpulen, die bei der Aufnahme an den Körper angelegt werden, wird keine Veränderung der Patientenlage angenommen, weshalb eine Transformationsmatrix nicht erforderlich ist. Die Bildabschnitte werden nach der Reihe nach (z.B. von Kopf bis Fuß) geladen und automatisch zusammengefügt. Angesichts der Intensitätsartefakte in den Übergangsbereichen zwischen den Segmenten (Abb. 1(a)), ist eine Korrektur dieser Regionen notwendig. Im ersten Schritt erfolgt eine Intensitätsnormalisierung. Die Helligkeit wird in diesem Bereich schematisch um einen Faktor erhöht, der sich aus den Mittelwerten und der maximalen Intensität der über- und unterhalb liegenden Bildschichten ergibt. Dabei wird die Intensität mittels eines Rescaling Filters auf das neue Maximum angepasst, welches sich nach $\hat{C}_{\max} = C_{\max} \cdot (\bar{R}/\bar{C})$ berechnet. \hat{C}_{\max} entspricht dem neuen, C_{\max} dem aktuellen Intensitätsmaximum. \bar{R} ist die mittlere Intensität des Referenzbereiches und \bar{C} die mittlere Intensität der zu korrigierenden Bildschicht. Die Aufnahmeartefakte auf der ersten sowie letzten Schicht jedes Segmentes werden durch eine Interpolation zwischen diesen und den Schichten über- und unterhalb unterdrückt. Zusammenfassend generieren die Methoden ein Ganzkörperdatensatz mit linearisierten Intensitätsübergängen zwischen den einzelnen Segmenten (Abb. 1(b)).

2.2 Korrektur der Dixon-Artefakte

Im Folgenden wird die Erkennung und Korrektur der Dixon-Artefakte beschrieben. Wie bereits in Kapitel 1 erläutert kann es bei dem Dixon MRT Protokoll zu einem Phasenfehler im Bild kommen. Aufgrund des Region Growing

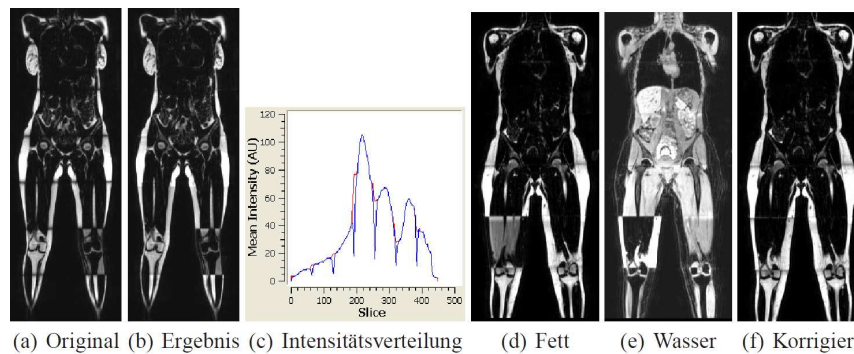


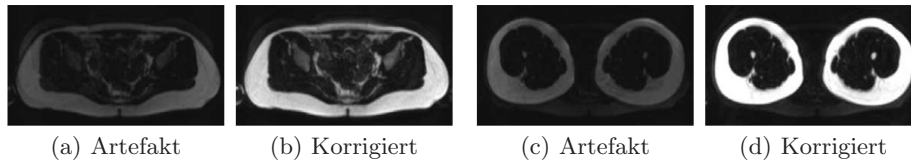
Abb. 1. Links: Generierung der korrigierten Ganzkörperdaten. Rechts: Korrektur der Artefakte in den unteren Extremitäten.

der Unwrapping-Methode tritt dieser Artefakt erfahrungsgemäß in den unteren Extremitäten auf. Daher liegt hier der Schwerpunkt auf der automatischen Erkennung und Korrektur. Im ersten Schritt erfolgt die Segmentierung und Binarisierung vom rechten und linken Bein, die im Folgenden voneinander getrennt bearbeitet werden. Diese getrennte Betrachtung ist notwendig, da das Artefakt meist nur in einem Bein auftritt (Abb. 1(a,d)). Um zu entscheiden welcher Abschnitt ein Artefakt enthält, wird ein Differenzbild zwischen den aneinandergrenzenden Randschichten der Segmente berechnet und anschließend das Volumen des Ergebnisbildes ermittelt. Diese Prozedur erfolgt in jeder Kombination zwischen den Wasser- und Fettdatensätzen. Ist das Volumen des Differenzbildes klein, sind die benachbarten Bildschichten ähnlich und gehören zu einer gemeinsamen Sequenz. Bei einem Artefakt ist das Volumen groß, da hier Informationen der Fettsequenz in einen Wasserdatensatz, oder umgekehrt, übergehen. Beginnend mit dem oberen Abschnitt im Bein, wird die Sequenz mit dem geringsten Volumenunterschied des Differenzbildes ausgewählt. Dies erfolgt in gleicher Weise für die weiteren Segmente, jeweils mit dem Ergebnis der vorherigen Region. Mit diesem Verfahren entsteht für das rechte und linke Bein jeweils eine Entscheidungs-Liste, an welcher Stelle die Fettsequenz durch die Informationen des Wasserdatensatzes ersetzt werden muss. Im letzten Schritt erfolgt die Ersetzung der Dixon-Artefakte durch die entsprechenden Daten der Wassersequenz, es entsteht ein korrigiertes artefakt-freies Ganzkörper-Fettbild (Abb. 1(f)).

3 Ergebnisse

Die Qualität der Bildschichten in der unteren und oberen Randregion jedes Segmentes wurde durch die Intensitätsnormalisierung optimiert. Abbildung 2(a) zeigt eine intensitätsschwache Bildschicht des Beckens, (b) die korrigierte Schicht nach Normalisierung der Intensität. Weiterhin wurde der Intensitätskontrast zwischen dem Objekt und den Aufnahmeartefakten durch die Interpolation vergrößert. In Abbildung 2(c,d) wird ein Beispiel für die Rauschunterdrückung gegeben. In der originalen Bildschicht hat das Aufnahmeartefakt einen ähnlichen Intensitätswert wie das Objekt. Im korrigierten Datensatz gibt es eine eindeutige Trennung zwischen diesen. Abbildung 1(a–c) zeigt ein Beispiel für das Gesamtergebnis des vorgestellten Verfahrens zur effizienten Generierung von Ganzkörperdaten. Es treten keine signifikanten Fehlanpassungen auf, was die Annahme der stabilen Patientenlage während der Aufnahme bestätigt. Weiterhin sind die Übergänge zwischen den Bildabschnitten geglättet und machen weitere Segmentierungen wie für die Fettgewebsanalyse möglich. Das Diagramm in Abb. 1(c) zeigt die Verteilung der mittleren Intensität jeder Bildschicht und demonstriert somit die Normalisierung der Übergangsbereiche. Die blaue Kurve repräsentiert das Original, die rote Linie das korrigierte Bild. Eine durchschnittliche Berechnungszeit von 10,13 Sekunden pro Ganzkörperdatensatz wurde mit einem Standard PC (Intel Core Quad 2,66GHz, 4GB RAM) erreicht. Damit stellt diese Methode einen schnellen Ansatz zur Generierung von Ganzkörperdaten dar. Die Korrektur der Dixon Sequenz erfolgte anhand von vier Probanden. Bei allen

Abb. 2. Aufbesserung der Bildrandbereiche. Links: Korrektur der intensitätsschwachen Bildschichten. Rechts: Korrektur der Aufnahmeartefakte.



wurde der Artefaktbereich innerhalb von $\varnothing 22,23$ s richtig erkannt und korrigiert. Abbildung 1(d-f) zeigt das Ergebnis der vorgestellten Methode. Bild 1d enthält den korrigierten Ganzkörperdatensatz mit dem Sequenzartefakt im rechten Oberschenkel. Bild 1e zeigt den Wasserdatensatz und Bild 1f das korrigierte artefaktfreie Ganzkörper-Fettbild.

4 Diskussion

In diesem Beitrag wurde ein neues Verfahren zur Generierung von Ganzkörperdaten präsentiert. Im Gegensatz zu anderen Verfahren ist der vorgestellte Ansatz aufgrund der nicht-überlappenden Eingabedaten unabhängig von Registrierungen. Die verrauschten Bildschichten in den Randregionen jedes Segmentes werden durch den kombinierten Algorithmus der Intensitätsnormalisierung und der Interpolation wiederhergestellt. Die Korrektur der Dixon-Artefakte wurde an allen Daten positiv getestet, jedoch werden weitere Daten benötigt um die Stabilität der Methode zu evaluieren. Einen Vergleich zu anderen Ansätzen kann nur schwer gezogen werden, da das Augenmerk in dieser Arbeit auf die Geschwindigkeit der Generierung und nicht auf die Qualitätsoptimierung von Ganzkörperdaten liegt. Unsere Methode ist jedoch durch die hohe Geschwindigkeit besser für die Segmentierungen und Diagnosen für die Fettgewebsanalyse bei großen Studien geeignet.

Literaturverzeichnis

1. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J Natl Cancer Inst.* 1981;66:1191–308.
2. Zhang C, et al. Abdominal obesity and the risk of all-cause, cardiovascular, and cancer mortality. *Circulation.* 2008 Apr;117:1658–67.
3. Lauenstein TC, et al. Whole-body MR imaging: evaluation of patients for metastases. *Radiology.* 2004;233:139–48.
4. Szumowski J, et al. Phase unwrapping in the three-point Dixon method for fat suppression MR imaging. *Radiology.* 1994; p. 555–61.
5. Glocker B, et al. MRI composing for whole body imaging. *Proc BVM.* 2009; p. 420–24.
6. Wolf I, et al. The medical imaging interaction toolkit (MITK). *Med Image Anal.* 2005; p. 594–604.

Iterative Closest Point Algorithm in the Presence of Anisotropic Noise

L. Maier-Hein, T. R. dos Santos, A. M. Franz, H.-P. Meinzer

German Cancer Research Center, Div. of Medical and Biological Informatics
l.maier-hein@dkfz-heidelberg.de

Abstract. The Iterative closest point (ICP) algorithm is a widely used method for 3D point set registration. It iteratively establishes point correspondences between two input data sets and computes a rigid transformation accordingly. From a statistical point of view, the algorithm implicitly assumes that the points are observed with isotropic Gaussian noise. In this paper, we present the first variant of the ICP which accounts for anisotropic localization uncertainty in *both* steps of the algorithm and show that in the presence of anisotropic noise, the modified ICP is better suited for both, establishment of point correspondences and transformation computation.

1 Introduction

The iterative closest point (ICP) algorithm [1] is a widely used method for 3D point set registration in the context of medical image processing. In order to find an optimal alignment between two input point clouds the algorithm iteratively (1) establishes point correspondences given the current alignment of the data and (2) and computes a rigid transformation accordingly. It can be shown that the method converges to an at least local minimum with respect to a mean-square distance metric. From a statistical point of view, the algorithm implicitly assumes that the points are observed with zero-mean, identical and isotropic Gaussian noise. In practice, however, point localization errors may be highly anisotropic. Optical tracking systems and time-of-flight (TOF) cameras, for instance, typically have a much higher localization uncertainty in the view direction of the camera. Although various ICP variants have been proposed in the literature [2], the issue of anisotropic noise has so far been given very little attention. This may be due to the fact that all known closed-form solutions for registering two point sets with known correspondences implicitly assume isotropic noise [3]. To our knowledge, Estépar [4] were the first and so far the only ones who proposed a variant of the ICP, the so-called generalized total least squares ICP algorithm (GTLS-ICP), that addresses this issue. The applied iterative method for computing a rigid transformation based on the current estimation of point correspondences accounts for anisotropic noise in both the target and the source models. However, the standard closest point operator based on the Euclidean distance is applied for establishing the point correspondences (Fig. 1(e)).

In this paper, we propose a closest point operator for handling point localization inaccuracies and show that in the presence of anisotropic noise, it is better suited for identifying corresponding points than the Euclidean distance. Furthermore, we present a variant of the ICP algorithm based on this operator and the recently introduced point registration algorithm by Balachandran and Fitzpatrick [3].

2 Materials and Methods

2.1 Closest Point Operator and ICP with Anisotropic Weighting

Let $P = \{\mathbf{p}_1, \dots, \mathbf{p}_{N_P}\}$ be a noisy source point set (e.g., a set of points acquired with a TOF camera) to be registered to a model point set $X = \{\mathbf{x}_1, \dots, \mathbf{x}_{N_X}\}$ (e.g., the set of vertices from a surface mesh extracted from computed tomography (CT) data). In the original ICP algorithm, the goal is to find the rigid body transformation that minimizes the mean squared Euclidean distance between corresponding points (Fig. 1(e)).

In order to account for anisotropic noise, we modify the original ICP algorithm as follows: In addition to the two point sets to be registered, a 3-by-3 weight matrix W_i is provided for each point \mathbf{p}_i in the source point set, reflecting the statistical noise model of the input data. Based on the definition of these weights W_i , the closest point operator is defined as follows:

$$C_{\text{new}}(\mathbf{p}, X) = \arg \min_{\mathbf{x}_i \in X} \|W_i(\mathbf{p} - \mathbf{x}_i)\|_2 \quad (1)$$

As in [3], anisotropic localization errors can be accounted for by giving less weight to the direction associated with the most noise (e.g., the axis along the direction of view of a TOF camera). In addition, outliers can be removed by setting $W_i = \mathbf{0}$. $W_i = I \forall i$ yields the standard version C_{original} of the closest point operator.

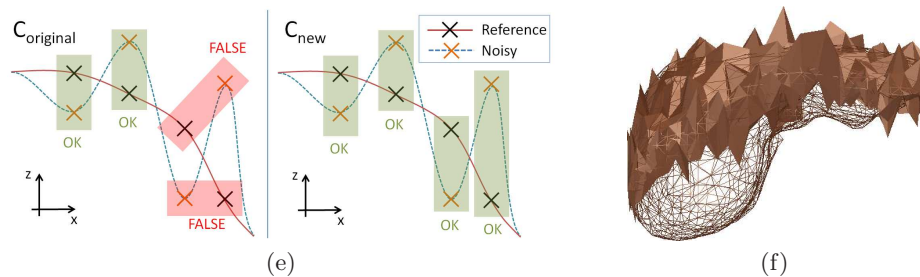


Fig. 1. (a) Schematic illustration of the establishment of point correspondences. Reference mesh (brown solid line), noisy mesh (dotted blue line) and correspondences (boxes) are shown for C_{original} (standard closest point operator) and C_{new} (new closest point operator with less weight given to the direction z). (b) Noisy submesh registered to a reference liver mesh via the proposed anisotropic ICP.

Given the weights W_i for each point in the noisy input data set, the aim of the anisotropic ICP algorithm is to find a rotation matrix R and a translation vector \mathbf{t} such that the following error metric is minimized

$$e(R, \mathbf{t}) = \frac{1}{N_P} \sum_{i=1}^{N_P} \|W_i (R\mathbf{y}_i + \mathbf{t} - \mathbf{p}_i)\|_2^2 \quad (2)$$

where $\mathbf{y}_i = C_{\text{new}}(\mathbf{p}_i, \hat{X})$ with $\hat{X} := \{R\mathbf{x}_j + \mathbf{t}\}, j = 1, \dots, N_X$. As in [3], we assume, that the localization error in each point is normally distributed with zero-mean and can be represented by three uncorrelated components along a set of orthogonal principal axes \mathbf{x} , \mathbf{y} and \mathbf{z} . We set $W_i = \text{diag}(\sigma_{ix}^{-1}, \sigma_{iy}^{-1}, \sigma_{iz}^{-1})$ with $\sigma_{ix}, \sigma_{iy}, \sigma_{iz}$ denoting the standard deviations associated with point \mathbf{p}_i .

The modified ICP works as follows:

1. Initialize variables: $k := 1; X^0 := X; e^0 := \infty; R_{\text{global}} := I; \mathbf{t}_{\text{global}} := \mathbf{0}$
2. Compute the current corresponding points $Y^k = \{\mathbf{y}_i^k\}, i = 1, \dots, N_P$ with $\mathbf{y}_i^k := C_{\text{new}}(\mathbf{p}_i, X^{k-1})$,
3. Compute the rotation matrix R^k ; the translation vector \mathbf{t}^k and the registration error e^k for mapping the point set Y^k onto the point set P using [3]

$$(R^k, \mathbf{t}^k) = \arg \min_{R, \mathbf{t}} \frac{1}{N_P} \sum_{i=1}^{N_P} \|W_i (R\mathbf{y}_i^k + \mathbf{t} - \mathbf{p}_i)\|_2^2 \quad (3)$$

$$e^k = \frac{1}{N_P} \sum_{i=1}^{N_P} \|W_i (R^k \mathbf{y}_i^k + \mathbf{t}^k - \mathbf{p}_i)\|_2^2 \quad (4)$$

4. Apply the rigid body transformation computed in the previous step to X^{k-1} to obtain the transformed model point set X^k .
5. Update the global rigid transformation

$$R_{\text{global}} = R^k R_{\text{global}}; \mathbf{t}_{\text{global}} = R^k \mathbf{t}_{\text{global}} + \mathbf{t}^k \quad (5)$$

6. if $|e^k - e^{k-1}| < \epsilon$ or the maximal number of iterations has been reached, terminate. Otherwise, set $k := k + 1$ and go to step 2.

Note that the algorithm maps the model point set to the noisy source point set because this allows us to perform all calculations in the coordinate system in which the noise model is given. The inverse transformation is given by the rotation matrix R_{global}^{-1} and the translation vector $-R_{\text{global}}^{-1} \mathbf{t}_{\text{global}}$.

We now show, that the presented algorithm converges to an at least local minimum with respect to the error metric e by proving the following theorem:

Provided that the iterative registration algorithm proposed in [3] yields the optimal rigid transformation with respect eq. 4, the registration error for mapping corresponding points decreases in each iteration, i.e., $e^k \leq e^{k-1} \forall k > 0$.

Proof by contradiction: Let us assume that $\exists k$ with $e^k > e^{k-1}$. Hence

$$\sum_{i=1}^{N_P} \|W_i (R^k \mathbf{y}_i^k + \mathbf{t}^k - \mathbf{p}_i)\|_2^2 > \sum_{i=1}^{N_P} \|W_i (R^{k-1} \mathbf{y}_i^{k-1} + \mathbf{t}^{k-1} - \mathbf{p}_i)\|_2^2 \quad (6)$$

Due to the definition of the closest neighbour (1), the following equation holds for each point in the noisy data set \mathbf{p}_i (note: $(R^{k-1}\mathbf{y}_i^{k-1} + \mathbf{t}^{k-1}) \in X^{k-1}$)

$$\|W_i(\mathbf{y}_i^k - \mathbf{p}_i)\|_2^2 \leq \|W_i(R^{k-1}\mathbf{y}_i^{k-1} + \mathbf{t}^{k-1} - \mathbf{p}_i)\|_2^2 \quad (7)$$

In consequence (due to $a_i \leq b_i \forall i \Rightarrow \sum a_i \leq \sum b_i$)

$$\sum_{i=1}^{N_P} \|W_i(\mathbf{y}_i^k - \mathbf{p}_i)\|_2^2 \leq \sum_{i=1}^{N_P} \|W_i(R^{k-1}\mathbf{y}_i^{k-1} + \mathbf{t}^{k-1} - \mathbf{p}_i)\|_2^2 \quad (8)$$

This, however, would mean that $\hat{R}^k := I$ and $\hat{\mathbf{t}} = 0$ would yield a better registration result than R^k and \mathbf{t}^k in step k of the algorithm (eq. 6), which is in contradiction to the presupposition of the theorem. Hence, $e^k \leq e^{k-1} \forall k$, q.e.d.

2.2 Experiments

To evaluate the proposed variant of the ICP algorithm, five liver meshes were extracted from human CT data (approx. 1000 vertices per mesh). For each liver, four submeshes were extracted representing a cranial, caudal, dorsal and ventral view on the liver respectively. Next, zero-mean, identical Gaussian noise with covariance matrices $\text{diag}(1, 1, \sigma^2)$ (cranial/caudal) or $\text{diag}(1, \sigma^2, 1)$ (dorsal/ventral) was added to the meshes, with $\sigma^2 \in \{1^2, 3^2, 5^2, 7^2, 9^2\}$ mm. Both, C_{original} and C_{new} were then applied to establish point correspondences between the two meshes, and the mean percentage of correct correspondences was determined. Next, both versions of the ICP algorithm were used to register the two point sets (Fig. 1(f)). After convergence, the percentage of correct point correspondences was computed again, and the translation error and the rotation

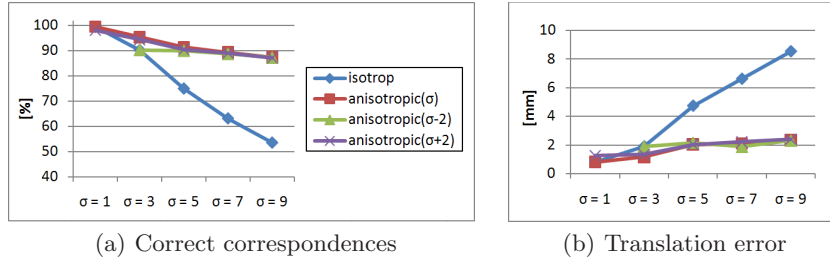


Fig. 2. Mean percentage of correct point correspondences before application of the ICP (a) and mean translation error after running the ICP (b) as a function of the standard deviation σ corresponding to the direction of the most noise (both averaged over 20 submeshes). *Isotropic* represents the standard closest point operator C_{original} and the standard ICP respectively. When C_{new} was applied, the component of the weight matrix corresponding to the most noise was set to $1/\sigma$, $1/(\sigma - 2)$ and $1/(\sigma + 2)$ for a correct noise estimation (anisotropic(σ)), underestimated noise (anisotropic($\sigma-2$)) and overestimated noise (anisotropic($\sigma+2$)) respectively, while the weights corresponding to the remaining two components was set to 1 (all values in mm).

error were determined as in [4], with the null vector and the identity matrix serving as ground truth.

3 Results

As shown in Fig. 2, the proposed closest point operator C_{new} performs considerably better than the standard operator in the presence of anisotropic noise. Even for a standard deviation of 9 mm along the view direction of the camera, the method yields a rate of correct correspondences of $87 \pm 5\%$ compared to $54 \pm 6\%$ with the standard method. Similar improvements were obtained for the translation error (2.3 ± 1.9 mm vs. 8.5 ± 7.9 mm) the percentage of correct correspondences *after* ICP convergence ($87 \pm 4\%$ vs. $52 \pm 6\%$) as well as for the rotation error ($0.2 \pm 0.1^\circ$ vs. $0.6 \pm 0.4^\circ$). Furthermore, the results indicate that the noise distribution must not be known exactly.

4 Discussion

In this paper, we presented the first variant of the ICP which accounts for anisotropic noise in *both* steps of the iterative algorithm: Computation of point correspondences and transformation computation. The presented results indicate that the proposed method is better suited for 3D point registration than the standard approach in the presence of anisotropic noise, even when the noise distribution is not exactly known. To establish the new method as a standard procedure, the following work remains to be done: (1) Efficient implementation for run-time optimization, (2) assessment of convergence speed and accuracy with random initial positions, (3) comparison with other variants of the ICP as well as with other methods for point registration in the presence of anisotropic noise (e.g. [4]), and (4) evaluation of the effects of anisotropic noise in the reference data set.

Acknowledgement. The authors would like to thank R. Balachandran and J. M. Fitzpatrick for their advice and for providing the source code of the algorithm.

References

1. Besl PJ, McKay ND. A method for registration of 3-D shapes. IEEE Trans Pattern Anal Mach Intell. 1992;14:239–56.
2. Rusinkiewicz S, Levoy M. Efficient variants of the ICP algorithm. In: Proc IEEE 3DIM; 2001. p. 145–12.
3. Balachandran R, Fitzpatrick JM. Iterative solution for rigid-body point-based registration with anisotropic weighting. Proc SPIE. 2009;7261:72613D.
4. Estépar RSJ, Brun A, Westin CF. Robust generalized total least squares iterative closest point registration. In: Proc MICCAI; 2004. p. 234–41.

Ein gradientenflussbasiertes Ähnlichkeitsmaß für das Tracking von Gefäßen

Xin Wang¹, Ivo Wolf¹, Philipp Hartmann¹, Tobias Heimann²,
Hans-Peter Meinzer¹, Ingmar Wegner¹

¹Abteilung für Medizinische und Biologische Informatik, DKFZ Heidelberg

²INRIA, Asclepios Projekt, Sophia Antipolis, Frankreich

`xin.wang@dkfz-heidelberg.de`

Kurzfassung. Die Segmentierung von Blutgefäßen ist essentiell für viele computergestützte medizinische Anwendungen. Trotz intensiver Forschung bleibt eine akkurate Segmentierung der Gefäßsegmente mit Stenosen und Kalzifikationen eine große Herausforderung. Dieser Beitrag beschreibt ein Ähnlichkeitsmaß für das Bayes'sche Tracking der Gefäßzentralinie. Dieser Ansatz erweitert ein gradientenflussbasiertes Ähnlichkeitsmaß um zusätzliche Informationen wie Grauwert- und Radiusänderungen entlang einer Trackingstrajektorie. Die Evaluierung wurde auf 7 CTA-Aufnahmen der Koronararterien mit Stenosen und Kalzifikationen durchgeführt. Die Ergebnisse zeigen, dass die Schätzungen der Gefäßverlaufsrichtung durch den erweiterten Ansatz im Durchschnitt um 22,57% verbessert wurden.

1 Einleitung

Die Segmentierung von Blutgefäßen aus volumetrischen Bilddaten ist wichtig für die Diagnose, die Therapie und die Operationsplanung. Darüber hinaus bildet die Segmentierung eine essentielle Basis für eine akkurate Visualisierung von Gefäßbäumen. Gefäße sind dünn, annähernd röhrenförmig und eingebettet in andere komplexe organische Strukturen. Zusätzlich können pathologische Veränderungen wie Stenosen auftreten. Eine Ursache dafür kann die Verkalkung sein. Auf den Bilddaten zeigen solche Gefäßsegmente nicht nur eine starke Verengung, sondern häufig auch einen großen Anstieg der Grauwerte. Dadurch stellt die Segmentierung der Gefäße im Allgemeinen eine große Herausforderung für die medizinische Bildverarbeitung dar. In diesem Themengebiet wurden bisher eine Vielzahl von Verfahren veröffentlicht [1]. Ein erfolgsversprechender Ansatz ist das Tracking der Gefäßzentralinie. Ein hierzu herangezogener Algorithmus ist das Partikelfilter [2]. Dieses Monte-Carlo-Verfahren schätzt rekursiv die posteriori Wahrscheinlichkeit eines dynamischen Prozesses. In der Bayes'schen Schätzung ist die Definition der Ähnlichkeit (engl. likelihood) ein zentraler Schritt. Die Ähnlichkeit beschreibt die Wahrscheinlichkeit, dass in der betrachteten Bildregion ein Gefäß ist. Das in [3] vorgestellte Ähnlichkeitsmaß basiert auf einer mathematischen Modellierung der Grauwertverteilung eines Gefäßes. In [4] wurde

ein Ähnlichkeitsmaß beschrieben, welches den nach innen gerichteten Gradientenfluss auf einer Kugeloberfläche misst. Beide Ansätze sind empfindlich gegen starke Grauwertänderungen innerhalb der Gefäße, die durch Kalzifikationen auftreten können. Basierend auf dem Ansatz in [4] wird ein Ähnlichkeitsmaß in dieser Arbeit vorgestellt, welches die Inhomogenität der Grauwertverteilung an den Stenosenstellen mit Kalzifikationen berücksichtigt.

2 Materialien und Methoden

2.1 Ein neues gradientenflussbasiertes Ähnlichkeitsmaß

Das in [4] beschriebene Ähnlichkeitsmaß für einen Bezugspunkt \mathbf{p} basiert auf einer Kugeloberfläche $K(\mathbf{p}, r)$ mit Radius r . Die Kugeloberfläche wird gleichmäßig in N Punkten \mathbf{x}_i diskretisiert. Das Ähnlichkeitsmaß wird definiert durch

$$F(\mathbf{p}, r) = \frac{1}{N} \sum_{i=1}^N \langle \nabla \mathbf{I}(\mathbf{i}), \mathbf{u}_i \rangle \quad (1)$$

Dabei bezeichnet $\nabla \mathbf{I}(\mathbf{i})$ den Gradienten an der Stelle \mathbf{x}_i und $\mathbf{u}_i = \frac{\mathbf{p} - \mathbf{x}_i}{|\mathbf{p} - \mathbf{x}_i|}$ die nach innen gerichteten radialen Richtungsvektoren. Während des Trackingsprozesses wird die Ähnlichkeit in jeder Iteration t an verschiedenen Positionen \mathbf{p}^t mit unterschiedlichen Radien r^t evaluiert. Die Position

$$\hat{\mathbf{p}}^t = \arg \max_{\mathbf{p}, r} F(\mathbf{p}^t, r^t) \quad (2)$$

wird benutzt, um die geschätzte Verlaufsrichtung der Gefäße

$$\mathbf{d}^t = \mathbf{p}^t - \mathbf{p}^{t-1} \quad (3)$$

aus der Position \mathbf{p}^{t-1} zu berechnen. Der geschätzte Radius an der Position \mathbf{p}^t ist

$$\hat{r}^t = \arg \max_{\mathbf{p}, r} F(\mathbf{p}^t, r^t) \quad (4)$$

In CTA-Datensätzen ist das Gefäßinnere normalerweise heller als das Gefäßäußere. Im idealen Fall ist $\hat{F}(\mathbf{p})$ maximal, wenn der Punkt \mathbf{p} auf der Zentrallinie des Gefäßes liegt und die Kugeloberfläche K die Gefäßwand schneidet. Aufgrund von Kalzifikationen im Gefäß können lokal sehr helle Regionen in der Nähe der Gefäßzentrallinie auftreten. In solchen Fällen ist $F(\mathbf{p}, r)$ maximal, wenn der Punkt \mathbf{p} innerhalb von der hellen Region liegt und die Kugeloberfläche K die Kanten dieser Region schneidet. Dies führt dazu, dass der Trackingsprozess die Richtung für den Weiterverlauf der Gefäße falsch einschätzt (Abb. 1 und 2). Die Bildregionen mit Kalzifikationen haben eine viel größere Intensität als das Gefäßinnere. Gleichzeitig ist der Umfang solcher Regionen häufig deutlich kleiner

Tabelle 1. Evaluierungsergebnisse der Ähnlichkeitsmaße F und F_{opt} auf 7 CTA-Aufnahmen der Koronararterien.

	$d(F_{\text{opt}})$	$d(F)$	Verbesserung
Daten 1	0.651	0.416	0.235
Daten 2	0.843	0.583	0.260
Daten 3	0.782	0.609	0.173
Daten 4	0.696	0.328	0.368
Daten 5	0.631	0.376	0.255
Daten 6	0.776	0.653	0.123
Daten 7	0.735	0.600	0.135

als der Gefäßradius. Aus diesen Beobachtungen wird eine neue Ähnlichkeit wie folgt definiert

$$F_{\text{opt}}(\mathbf{p}^t, r^t) = \left(1 - \frac{|\hat{r}^{t-1} - r^t|}{\hat{r}^{t-1} + r^t}\right) (F(\mathbf{p}, r) - \alpha |g^{t-1}(\mathbf{p}^{t-1}, r^{t-1}) - g^t(\mathbf{p}^t, r^t)|) \quad (5)$$

Dabei bezeichnet t den Iterationsschritt, α den Gewichtungsfaktor, und $g(p, r)$ den mittleren Grauwert innerhalb der betrachteten Kugel K . Durch die Terme $\left(1 - \frac{|\hat{r}^{t-1} - r^t|}{\hat{r}^{t-1} + r^t}\right)$ und $-\alpha |g^{t-1}(\mathbf{p}^{t-1}, r^{t-1}) - g^t(\mathbf{p}^t, r^t)|$ werden jeweils große Radius- und Grauwertänderungen der Kugel K zwischen zwei nacheinander folgenden Iterationen bestraft. Dadurch können die oben genannten Schätzungsfehler reduziert werden.

2.2 Evaluierung

Zur Evaluierung werden CTA-Datensätze von Koronararterien des MICCAI-Workshops 2008 [5] verwendet. Mit Hilfe der bereitgestellten Referenzpositionen (engl. ground truth) wurden 7 Stenosen mit Schwellwertverfahren semiautomatisch klassifiziert. Zum quantitativen Vergleich der beschriebenen Ähnlichkeitsmaße wird ein Maß wie folgt definiert

$$d = \frac{n_r}{n_{\text{all}}} \quad (6)$$

Dabei bezeichnet n_{all} die gesamte Anzahl der Referenzpositionen im betrachteten Gefäßsegment. Weiter stellt n_r die Anzahl der Referenzpositionen dar, an denen die Richtungsschätzung anhand des Ähnlichkeitsmaßes als richtig klassifiziert wurde. Die geschätzte Richtung wird als korrekt betrachtet, wenn die nächste Suchposition entlang der geschätzten Richtung mit Schrittweite \hat{r} innerhalb des Toleranzbereiches Q liegt. Der Toleranzbereich Q kann als ein synthetisches Gefäß angenommen werden, welches die gleiche Zentrallinie, aber nur die Hälfte des Radius des zu segmentierenden Gefäßes hat.

Tabelle 2. Evaluierungsergebnisse der Ähnlichkeitsmaße F und F_{opt} auf 7 extrahierten Gefäßsegmenten mit Stenosen und Kalzifikationen.

	$d(F_{\text{opt}})$	$d(F)$	Verbesserung
Stenosen 1	0.686	0.438	0.248
Stenosen 2	1.000	0.818	0.182
Stenosen 3	0.615	0.385	0.230
Stenosen 4	0.667	0.583	0.084
Stenosen 5	0.524	0.311	0.213
Stenosen 6	0.350	0.000	0.350
Stenosen 7	0.455	0.182	0.273

3 Ergebnisse

Die Evaluierungsergebnisse der Ähnlichkeitsmaße F und F_{opt} werden in den Tab. (1) und (2) dargestellt. In Tab. (1) werden die beiden Ähnlichkeitsmaße anhand von 7 CTA-Aufnahmen der Koronararterien miteinander verglichen. Um den Effekt beider Ansätze an Stenosenstellen mit Kalzifikationen zu zeigen, fokussiert sich der Vergleich in Tab. (2) auf 7 Stenosensegmente, die jeweils aus einem Gefäß extrahiert wurden. Abb. (1) stellt den Beispieldatensatz einer Koronararterie dar. Im Vordergrund ist eine leichte Verengung der Gefäße und eine sehr helle Region aufgrund der Kalzifikation erkennbar. In Abb. (2) werden die geschätzten Verlaufsrichtungen der Gefäße visualisiert, die jeweils anhand der Ähnlichkeitsmaße F (rot) und F_{opt} (grün) berechnet wurden. Das veranschaulichte Gefäßsegment wurde aus dem Datensatz in Abb. (1) extrahiert.

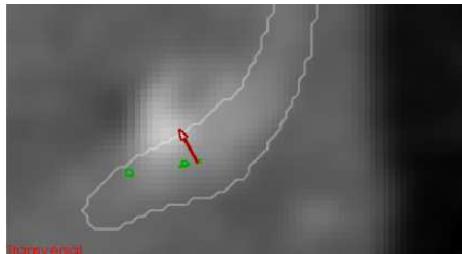
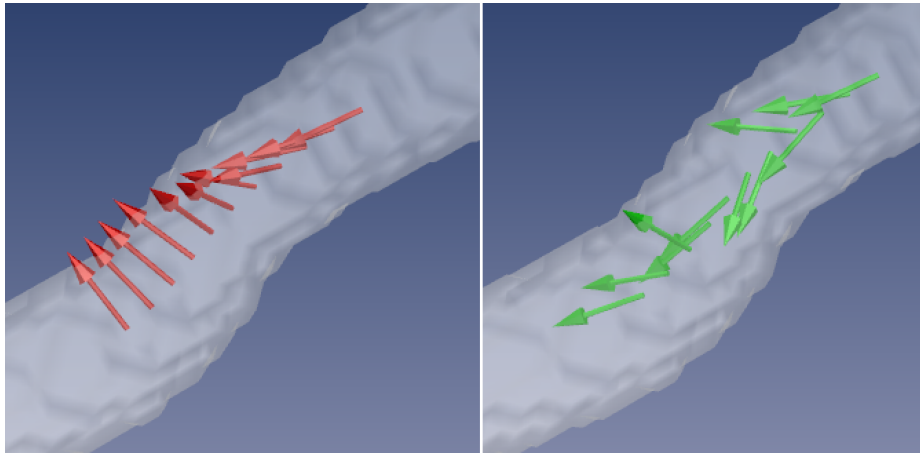


Abb. 1. Ansicht des CTA-Datensatzes einer Koronararterie auf einer transversalen Ebene. Die weiße Linie visualisiert die äußere Kontur des Gefäßes. An diesem Gefäßsegment ist eine leichte Verengung mit einer sehr hellen Region erkennbar. Die roten und grünen Pfeile stellen jeweils die geschätzten Verlaufsrichtungen anhand der Ähnlichkeitsmaße F und F_{opt} dar.

4 Diskussion

In diesem Beitrag wird ein Ähnlichkeitsmaß für das Bayes'sche Tracking von Blutgefäßen beschrieben. Dieser Ansatz erweitert ein gradientenflussbasiertes

Abb. 2. 3D-Visualisierungen der geschätzten Verlaufsrichtungen von Gefäßen anhand der Ähnlichkeitsmaße F (links) und F_{opt} (rechts). Das durch die weiße Oberfläche dargestellte Gefäßsegment zeigt eine leichte Stenose. Dieses Gefäßsegment wurde aus dem Datensatz in (Abb. 1) extrahiert.



Ähnlichkeitsmaß um zusätzliche Informationen wie Grauwert- und Radiusänderungen entlang der Trackingstrajektorie. Das Verfahren wurde anhand von 7 CTA-Datensätzen von Koronararterien evaluiert. Die Ergebnisse zeigen, dass die Schätzungsgenauigkeit für die Verlaufsrichtungen der Gefäße durch die Erweiterung des Ähnlichkeitsmaßes im Allgemeinen erhöht wurde. Ein besonderer Vorteil des neuen Ähnlichkeitsmaßes F_{opt} zeigt sich an den Gefäßsegmenten mit Stenosen und Kalzifikationen. Die durchschnittliche Verbesserung an solchen Stellen beträgt 22,57%. Künftige Arbeiten beinhalten zum einen die Optimierung des Parameters α in (5) und zum anderen die Optimierung der Laufzeit bei der Auswertung der Ähnlichkeit.

Literaturverzeichnis

1. Lesage D, Angelini ED, Bloch I, et al. A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes. *Med Image Anal.* 2009;in press.
2. Schaap M, Manniesing R, Smal I, et al. Bayesian tracking of tubular structures and its application to carotid arteries in CTA. *Proc MICCAI.* 2007;10:562–70.
3. Worz S, Rohr K. Segmentation and quantification of human vessels using a 3-D cylindrical intensity model. *IEEE Trans Image Process.* 2007;16(8):1994–2004.
4. Lesage. Medial-Based bayesian tracking for vascular segmentation: Application to coronary arteries in 3D CT angiography. *Int J Biomed Imaging.* 2008; p. 268–71.
5. Metz CT, Schaap M, van Walsum T, et al. 3D Segmentation in the clinic: A grand challenge II: coronary artery tracking. *Insight.* 2008.

Reconstruction Image Quality Theory

Evaluation of Filtered Back-Projection and Ordered-Subset Expectation Maximization

Herfried Wiecezorek

Philips Research Laboratories, Aachen, Germany
`herfried.wieczorek@philips.com`

Abstract. We have developed an image quality theory for reconstruction that we apply to filtered back-projection (FBP) and statistical reconstruction (OSEM) for Single Photon Emission Computed Tomography (SPECT). Quantitative measures of reconstruction performance are given in terms of signal and noise power spectra, SPS and NPS, that we derive from phantom images. This allows evaluating the properties of statistical reconstruction, especially signal recovery, noise, impact of phantom size, and detector resolution. Our analysis shows how noise in reconstructed images is reduced by iterative resolution recovery.

1 Introduction

Single Photon Emission Computed Tomography (SPECT) is a standard imaging modality in nuclear medicine, especially in the cardiology segment. SPECT images provide the basis for clinical diagnosis of coronary artery disease, and combined with an attenuation map provided e.g. by computed tomography, the evaluation of absolute tracer uptake will be possible. Necessary pre-conditions for quantitative imaging are a model of sensitivity and spatial resolution of SPECT detectors and the understanding of reconstruction properties [1]. In this paper we present an assessment of reconstruction image quality based on an analysis of filtered back-projection, an analytic reconstruction method that is still standard for computed tomography [2]. For SPECT imaging, iterative methods like MLEM [3] and OSEM [4] are preferred for their noise performance and the possibility to correct for attenuation, spatial resolution, scatter, and cross-contamination of different radioisotopes [5]. We investigate FBP and OSEM using software phantoms and evaluate the performance of these reconstruction methods in terms of signal and power spectra, a method known as the Detective Quantum Efficiency (DQE) model from x-ray imaging.

2 Materials and Methods

Description of image quality is based on a contrast phantom, a cylinder of radius r_b with a spherical lesion of radius r_l in its centre. Activity is measured by a

SPECT detector with parallel hole collimator, characterized by square detector pixels of pitch p and detector efficiency E , defined by NEMA system planar sensitivity [6].

The SPECT image is reconstructed from projections taken under different viewing angles equally spread over 360° during the imaging time T .

Activity per voxel is given in units of Becquerel, specified as A_l and A_b in the lesion and background region, respectively. When we take a transaxial slice of thickness p and measure the radiation from the phantom slice with a detector in close contact, the average number of background counts per pixel in the detector is $n_b = A_b T E$, and the signal-to-noise ratio is $SNR_0 = \sqrt{A_b T E}$.

Lesion detectability is described by the Rose criterion, stating that the contrast-to-noise ratio $CNR = \Delta S/N$ has to be larger than a value $k \approx 5$ [7]. With the lesion contrast C_0 we get $CNR_0 = C_0 \cdot SNR_0 \cdot \sqrt{r_l^2 \pi / p^2}$. We use a software phantom with $r_b = 75\text{mm}$, $r_l = 25\text{mm}$, voxel size $p = 3\text{mm}$ extend our theory to statistical reconstruction. To speed up simulation we calculate the central transaxial slice of the phantom in 2D cylinder geometry. With a background activity of 70MBq per litre, a square hole collimator with $E = 1.99 \cdot 10^{-4}$ ($L = 30\text{mm}$ septa length, $D = 1.5\text{mm}$ hole size, infinite septa absorption at negligible septa thickness), 128 projection angles and 20s time per view we calculate a reconstructed signal value of $A_b T E = 995$ counts per voxel.

Detector resolution is considered analytically in our simulation. When the collimator is 5mm away from the phantom surface, spatial resolution in the phantom centre is calculated as $R = 5.5\text{mm}$ parallel to the detector axes, equivalent to a Gaussian distribution with $\sigma = R/2p\sqrt{2 \ln 2} = 0.78$ voxels.

Noise Power Spectra are extracted by Fourier analysis of the centre part of the reconstructed phantom without lesion. We use a 16×16 pixel region, apply a one-dimensional Fourier transform on 16 rows with their respective mean values subtracted, and take the average of the squared moduli to generate the noise power spectra. All spectra are symmetrically defined in the frequency range $-0.5 \dots 0.5$ in units of cycles per pixel. By definition, the integral of the NPS function is equal to the background variance of the image.

Signal Power Spectra are derived by Fourier transformation of a differentiated step function. We differentiate the edge of the lesion with contrast $C_0 = 1$ in a noise free image along one of the detector axes, apply a one-dimensional Fourier transform on 16 data values centred on the lesion edge, and calculate the signal power spectra from the squared moduli. The zero frequency value of the SPS function is equal to the step height at the lesion edge which is, for ideal SPECT, given by $SPS(0) = C_0 \cdot A_b T E$. The normalized SPS function is better known as the squared modulation transfer function, MTF^2 .

3 Results

In the following we start with the evaluation of FBP and OSEM for ideal projections, and then add corrections for attenuation and detector resolution in statistical reconstruction.

3.1 Ideal SPECT

In filtered back-projection the signal of all detector pixels is added up and spatially filtered so that the value for the number of counts in a central voxel is restored [2]. Noise is added on all voxels along the lines of back-projection, resulting in an inverse square root dependence on the number of pixels per phantom diameter, $SNR = SNR_0 \sqrt{p/2r_bQ}$, with Q describing the effect of the filter function [1]. The same holds for the contrast-to-noise ratio CNR_0 . Simulated data show perfect agreement with this noise theory.

Fig. 1 shows noise power spectra for ideal SPECT reconstructed with FBP and linear interpolation to avoid streaking, in comparison to OSEM reconstruction, both shown smoothed with B-splines. For Ram-Lak and Shepp-Logan filters we find a nearly white spectrum at lower frequency and moderate noise reduction at higher frequency due to linear interpolation. Cosine and Hann filters suppress noise strongly especially in the high frequency range, giving a different noise texture. Additionally, noise is reduced throughout the whole frequency range.

Iterative reconstruction is known for soft, low-noise images at low iteration numbers and for sharper images and higher noise at a high number of iterations. For OSEM with low iteration numbers we see that noise power spectra decrease exponentially by nearly one order of magnitude from zero to 0.5 cycles per pixel. Duplication of the iteration numbers shifts up the noise power spectra by a factor of two or three, with a slight saturation in the lower frequency part. After 16 iterations the noise spectrum is approximately white, with somewhat lower values in the low and high frequency range.

Signal power spectra for ideal SPECT show the limited recovery of the high frequency signal due to linear interpolation and different filters, most prominent for the Hann filter (Fig. 2). Taking the measured curve for Ram-Lak as a reference for the impact of linear interpolation and multiplying by the squared theoretical filter functions we get the SPS curves for FBP with linear interpolation. The resulting data points (open symbols) show an excellent agreement with the measured signal power spectra.

With OSEM, after one iteration only the lowest frequency part of the spectrum is entirely recovered while the high frequency part is completely lost. With each doubling of iteration numbers the spectra are shifted up roughly by a factor of three before they saturate at the theoretically achievable signal level. With too high iteration numbers artifacts and artificial structures appear in the SPS.

3.2 Phantom Size Dependence

Fig. 3 shows SPS and NPS for two phantoms of different size in comparison. For the larger phantom, exactly the doubled number of iterations is needed to get the same SPS spectra as for the standard phantom, and the resulting noise power spectra are two times higher than for the smaller phantom. This is exactly the same phantom size dependence as for filtered back-projection.

3.3 Spatial Resolution

Correction for detector resolution is implemented in iterative reconstruction by application of a Gaussian filter in forward projection. We have to apply slightly smaller σ values than calculated to avoid overshoots in the power spectra. Fig. 4

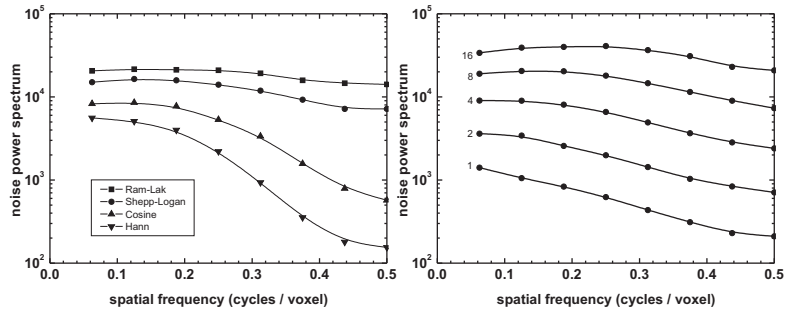


Fig. 1. Noise power spectra for ideal SPECT, FBP with linear interpolation and different filters (left), and OSEM, subset size 16, with number of iterations (right).

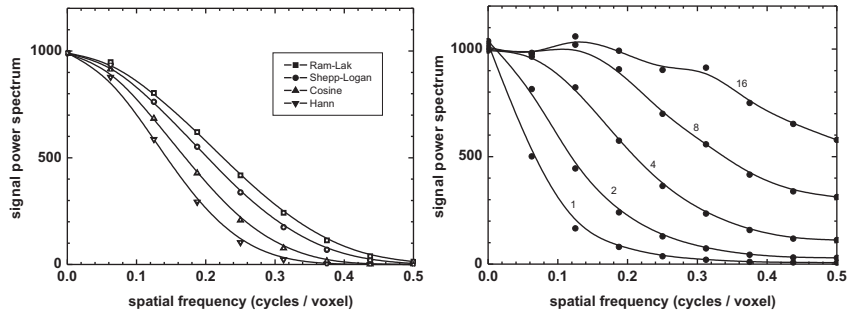


Fig. 2. Signal power spectra for FBP (left) and OSEM (right) as in Fig. 1.

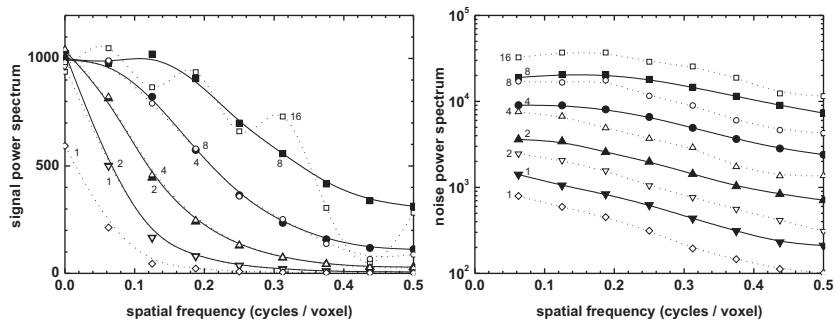
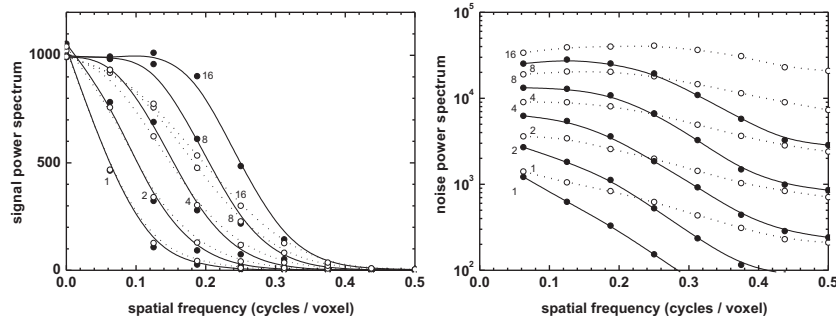


Fig. 3. Signal and noise power spectra for OSEM, subset size 16. Standard phantom (solid lines, full symbols) and 2 x larger phantom (dotted lines, open symbols).

Fig. 4. Signal (left) and noise power spectra (right) for OSEM with resolution recovery (solid lines) in comparison to uncorrected data (dotted lines).



shows the improved SPS for resolution correction with $\sigma = 0.65$ (solid lines) in comparison to uncorrected data (dotted lines). With a sufficient number of iterations we see that higher frequency signals can be recovered. The method is however limited by reconstruction time and increased noise. In the corresponding noise power spectra we see a general noise reduction, compared to uncorrected data, caused by resolution recovery, especially in the high frequency range. The variance of the image in this example is reduced by a factor of two.

4 Discussion

We present a theory describing the impact of reconstruction on image quality, providing quantitative measures for the assessment of reconstruction methods. We find a comparable behaviour of FBP and OSEM with reference to the object size. Directions are given how to optimize image quality in nuclear medicine by OSEM with the right choice of iteration numbers and use of resolution recovery. Future work will determine the limits of absolute quantification of tracer uptake.

References

1. Wiecek H, Goedicke A. Analytical model for spect detector concepts. *IEEE Trans Nucl Sci.* 2006;53(3):1102–12.
2. Kak A, Slaney M. *Principles of Computerized Tomographic Imaging.* New York, IEEE Press; 1987.
3. Shepp L, Vardi Y. Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging.* 1982;1(2):113–22.
4. Hudson, H M , Larkin, R S. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging.* 1994;13(4):601–9.
5. Dey, T , Wiecek, H , Backus, B , et al. Thallium-Stress, Technetium-Rest Protokoll für Cardiac SPECT. *Proc BVM.* 2010;in press.
6. NEMA standards publication NU1-2001. Rosslyn, VA USA, National Electrical Manufacturers Association; 2001.
7. Rose A. *Advances in Electronics.* Academic, New York; 1948.

MITK-DI

A new Diffusion Imaging Component for MITK

Klaus Fritzsche, Hans-Peter Meinzer

Division of Medical and Biological Informatics, DKFZ Heidelberg
k.fritzsche@dkfz-heidelberg.de

Abstract. Diffusion-MRI provides a unique window on brain anatomy and insights into aspects of brain structure in living humans that could not be studied previously. There is a major effort in this rapidly evolving field of research to develop algorithms that provide detailed information on the white matter fibre architecture and disorders in the brain. High angular resolution diffusion imaging (HARDI) provides an extraordinarily high level of detail, including entire manifolds of information at each voxel. The lack of standardized software tools for data I/O, model reconstruction, interactive visualization, feature extraction and statistics impedes development and sustainable evaluation in HARDI research. In this paper, a new toolkit component of the Medical Imaging and Interaction Toolkit (MITK), MITK-DI, is presented and its features are summarized at each stage of the pipeline from raw data processing to comprehensive statistics. MITK-DI aims at integrating novel diffusion imaging techniques into the MITK toolkit to provide a comprehensive software framework for performant data processing, analysis and interactive visualization. An exemplary application of the framework is shown using phantom and *in-vivo* datasets from a clinical MR scanner. Visualizations are presented that reveal details for which there would be no direct way to appreciate using conventional display of scalar- or RGB-images.

1 Introduction

Diffusion-MRI provides a unique and sensitive probe for the architecture of biological tissues. The diffusion tensor model, which was introduced in 1994 [1], encapsulates the measured diffusion properties of water molecules as the 3x3 covariance matrix of a Gaussian distribution. The well known limitations of this model in complex fiber configurations have led to the development of high angular resolution diffusion imaging (HARDI) techniques like q-ball imaging (QBI) [2]. Here, the information in each voxel is represented by a manifold, called orientation distribution function (ODF), that resembles the marginal displacement probability of diffusing water molecules after being released at the center of each voxel. While QBI has several advantages over diffusion tensor imaging (DTI), it significantly raises the complexity of data processing and visualization.

Several software toolkits like MedINRIA,¹ Camino,² SCIRun,³ Slicer⁴ or TrackVis⁵ support processing and visualization of diffusion imaging data. However, display and interaction capabilities are often limited. Camino, for example, uses unix-style command interfaces and manual export of models for viewing in other programs. There is a clear need for software tools that feature proper interactive visualization and quantitative evaluation in HARDI data.

In this paper, we present a new toolkit component of the Medical Imaging and Interaction Framework (MITK) [3], called MITK Diffusion Imaging (MITK-DI). MITK-DI aims at supporting cutting edge diffusion imaging techniques, extending the MITK framework in terms of data I/O, processing and visualization of diffusion related images. It will be made publicly available under a BSD-style open-source license in December 2009. As an example of application, several phantom and *in-vivo* datasets were processed and visualized in order to underline the capabilities of the component. In contrast to most other frameworks, MITK-DI addresses all aspects of application design including full integration into an OSGi-based⁶ application platform and fluent workflows. While other diffusion imaging toolkits exclusively focus on providing diffusion related functionalities, MITK-DI is tightly integrated into the grown-up platform MITK and therefore allows covering the complete cycle from raw-data to computer-aided diagnosis and statistics.

2 Materials and Methods

Standardization is a key issue in medical imaging research and can only be successful with a consistent and flexible software design, the publication of source code and the use of development tools that allow integration in a clinical environment with reasonable effort to prove relevance and, finally, to improve health care. Specific care was taken to these requirements during the design of MITK-DI.

2.1 Software Design and Application-level Support

Like the Visualization Toolkit (VTK⁷), the Insight Toolkit (ITK⁸) and MITK, MITK-DI is an object-oriented, cross-platform component implemented in C++. Most classes are derived from top-level classes of ITK, reusing all smart-pointer-, time-stamp-, pipeline- and parallel processing mechanisms. The

¹ <http://www-sop.inria.fr/asclepios/software/MedINRIA/>

² <http://www.cs.ucl.ac.uk/research/medic/camino/>

³ <http://www.sci.utah.edu/>

⁴ <http://www.slicer.org/>

⁵ <http://www.trackvis.org/>

⁶ <http://www.osgi.org/>

⁷ <http://www.vtk.org>

⁸ <http://www.itk.org/>

component extends MITK using the provided module-mechanism. Application-level classes are implemented as bundles for the BlueBerry application platform (<http://www.blueberry-project.org/>) that is the basis for all current MITK applications. Applications run on Windows, Mac OS X and Linux with native look and feel and 64-bit support. All MITK-DI bundles feature an intuitive GUI front-end with the support for multiple dataset processing.

2.2 Datastructures

MITK-DI includes algorithms for direct DICOM import and reading and writing of the widely used NRRD^{9,10} format. MITK-DI defines new classes that represent diffusion weighted images (`mitk::DiffusionImage`), q-ball images (`mitk::QBallImage`) and tensor images (`mitk::TensorImage`). All these data types make use of the MITK factory mechanism and can be handled via drag and drop on the application level using the file extensions `.dwi`, `.qbi` and `.dti` respectively.

2.3 Model Reconstruction

MITK-DI supports a range of standard and advanced reconstruction algorithms. It extends the opportunities given by ITK and allows for linear least squares, nonlinear least squares, weighted least squares and maximum likelihood fitting of diffusion tensors making use of the teem library.¹¹ It also allows for numerical q-ball reconstruction of HARDI data sets [4].

2.4 Visualization

A core feature of MITK-DI is the possibility to interactively visualize tensors and ODFs in multiple planar views. A mapper for concurrent visualization of colormaps and ODFs has been implemented that allows on the fly generation of the scene avoiding to hold the whole image representation in memory. A level-of-detail (LoD) mechanism allows for fluent navigation in the datasets. Different options for normalization and scaling can be configured during runtime. Rendering of vector images (e.g. principal diffusion direction) and the raw diffusion signal is possible.

2.5 Image Features and Statistics

Standard scalar measures such as the tensor's fractional anisotropy or relative anisotropy or the ODF's generalized fractional anisotropy can be computed. MITK's interactive segmentation framework and the image statistics bundle provide powerful tools for performing statistics on populations of voxels in certain ROIs.

⁹ <http://teem.sourceforge.net/nrrd/format.html>

¹⁰ http://wiki.na-mic.org/Wiki/index.php/NAMIC_Wiki:DTI:Nrrd_format

¹¹ <http://teem.sourceforge.net/>

2.6 Image Data and Experiments

Diffusion weighted MR images were acquired on a 3.0 T MR scanner (Magnetom TRIO, Siemens Medical Solutions, Erlangen, Germany) using a single shot, spin echo, echo planar imaging sequence. Phantoms were constructed in analogy to the production process described in [5]. The phantom imaging parameters were: FOV = 200×200 mm², in plane resolution = 1×1 mm², TR = 3.4 s, TE = 118 ms, 5 averages, $b = 3500$ s/mm², 252 gradient directions. *In-vivo* imaging parameters were: FOV = 240×205 mm², resolution = $2.5 \times 2.5 \times 2.5$ mm³, TR = 5 s, TE = 120 ms, 2 averages, $b = 3500$ s/mm², 64 gradient directions.

2.7 Availability

The latest version of the source code is available from the MITK web site¹².

3 Results

A Performance analysis was performed on a Intel Core2 Quad Processor (Q6600 at 2.40 GHz each) with 4 Gb of RAM. Q-ball reconstruction for ODFs with 252 directions took 12.91 s (image dimension 96x82x40, 64 directions). A linear diffusion tensor fit took 2.81 s for the same volume. Fluent interaction was achieved by the LoD-mechanism even for highly detailed and large datasets. Fig. 1 and Fig. 2 show renderings of the phantom acquisition and the *in-vivo* dataset respectively. It can be seen that q-ball ODFs provide significantly more detail than the tensor ODFs especially in crossing regions.

4 Discussion

We have introduced a new set of tools for processing and interactive visualization of diffusion MRI datasets. MITK-DI implements DTI and HARDI techniques

¹² <http://www.mitk.org/>

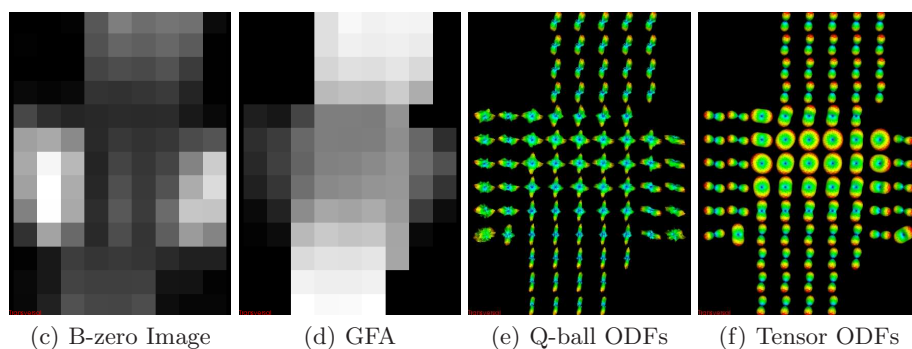
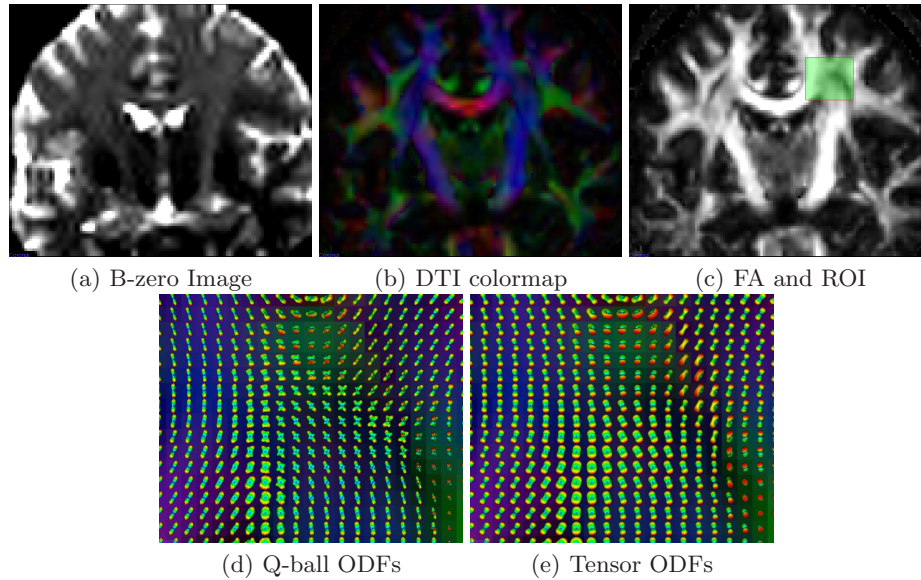


Fig. 1. Visualized HARDI acquisition of the 90° crossing fiber phantom.

Fig. 2. Coronal slice of the *in-vivo* HARDI acquisition with a detailed view on the crossing of corpus callosum and cortico-spinal tract, which is highlighted in 3(c).



and features display of intricate details in combination with fluent interaction. We will be applying these visualization tools to understanding aspects of HARDI imaging to examine disease characteristics. The integration of fibertracking algorithms and further reconstruction schemes is currently being worked on. Although primarily used as a research tool, diffusion weighted imaging is starting to find it's way to clinical application. MITK-DI is an attempt to bundle up and standardize current techniques in the field in a single framework and, following the open-source spirit, to enable other researchers to contribute and build upon it. Participation is welcome.

References

1. Basser PJ, Mattiello J, LeBihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. *J Magn Reson B*. 1994;103(3):247–54.
2. Tuch DS. Q-ball imaging. *Magn Reson Med*. 2004;52(6):1358–72.
3. Wolf I, Vetter M, Wegner I, et al. The medical imaging interaction toolkit. *Med Image Anal*. 2005;9(6):594–604.
4. Fritzsche KH, Meinzer HP. Numerical q-ball image reconstruction: an itk implementation. *Insight J*. 2009; p. 643.
5. Laun FB, Huff S, Stieltjes B. On the effects of dephasing due to local gradients in diffusion tensor imaging experiments. *J Magn Reson Imaging*. 2009;27(4):541–8.

MEDOX

An XML-based Approach of Medical Data Organization for Segmentation and Simulation

Eduard Fried¹, Yan Geng¹, Sebastian Ullrich², Daniel Kneer², Oliver Grottke³,
Rolf Rossaint³, Thomas M. Deserno¹, Torsten Kuhlen²

¹Department of Medical Informatics, RWTH Aachen University

²VR Group, RWTH Aachen University

³Department of Anaesthesia, University Hospital Aachen

efried@mi.rwth-aachen.de

Abstract. Applications in simulation or visualization often rely on data from multiple sources (e.g., MRI scans, extracted anatomical objects in a 3D representation). However, structured data organization is either non-existent (e.g., simply stored in folders), proprietary or specific (e.g., DICOM archives). In this paper, a novel extensible approach to organize and access arbitrary medical data is proposed, called Medical Data Organization with XML (MEDOX). The concept is based on a lightweight XML database, complementary to recent MedX3D standard developments. Data files are stored in a file system and referenced along with additional meta data in an XML file. Basic versioning of data objects is also supported. A simple C/C++ library was implemented to abstract the access patterns to the database and to avoid the error-prone manual manipulation of the XML file. As a proof of concept, it has been applied to regional anesthesia simulation in virtual environments.

1 Introduction

In a project where subject specific datasets are used for segmentation, simulation and visualization, the data has to be organized in an efficient way. Usually, it is sufficient to work with the newest data objects, yet sometimes access to all intermediate files in various versions is required. The organization of data objects has to support the user in finding and selecting the required data.

The simple storage of data files in the file system without additional meta information is error-prone and also complicates data exchange as it requires explicit user interaction and knowledge of the locations of all files.

The storage of all data in a relational database (e.g., SQL-based [1]) hides the storage details and requires additional tools to access the data. A human readable representation of the data is the X3D format [2], which is based on the XML standard. This data format supports the organization of a 3D scene, including the definition of the 3D-objects itself. Additional data, like DICOM files, is not directly supported. There is a working group, MedX3D [3], that is developing new standards for medical usage. However, the work is still in

early drafts and mainly focuses on volume rendering. The Foundational Model of Anatomy (FMA [4]) defines an ontology for anatomical structures, but it provides no means to organize or reference data files. Working with versioned files is a core functionality of tools like Subversion [5]. They offer an efficient storage model (by saving differences between versions) and also provide operations for comparison and merging. The primary requirement in this case would be the version history of data and configuration files, though.

For a simple and efficient usage of the available data, the database has to meet certain criteria. As every data object results from a processing (or imaging) operation, it should be linked to this process in a unique way. Consequently, the processing model should be represented in the data organization scheme. Data access should not distract the user from working with the files. Ideally, the access operations can be directly integrated into the processing tools.

2 Materials and Methods

In this paper, the authors focus on the usage of subject specific datasets, which initially consist of imaging data and also contain data from subsequent processing steps. This data originates from medical imaging, segmentation algorithms etc. and it can be used for visualization, simulation or analysis purposes (Fig. 1). The data for each subject is structured by a database file based on a novel XML-specification described in the following paragraphs.

The basic structure of the XML database contains meta data about the patient, the data directory and a set of data nodes (Fig. 2). The data directory represents the application specific data organisation and its dependencies (Fig. 2). To allow maximum flexibility, it is stored as a hierarchical tree structure with `<element>` tags to declare data objects and `<structure>` tags to group elements and other structures. The element tag uses the `id` attribute to link the declaration to the data object definition and the `name` and `description` attributes to give the object an application specific name.

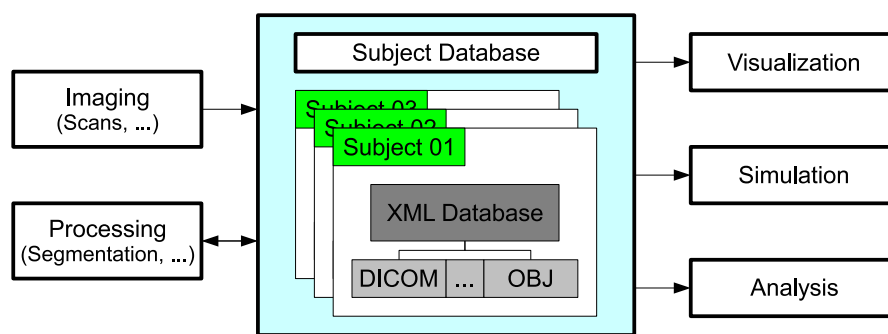


Fig. 1. Database architecture overview.

Inside the data node, there is a further distinction between image data and 3D-object data. The images can be raw data (e.g., medical scans) or processed images (e.g., denoising, segmentation). It is left to the user to extend the definition of the `<processed>` node with application specific processing parameters. As imaging data covers only a specific body region, this information is used to separate the data nodes.

The 3D-objects are referenced inside a set of segmentation nodes. As these objects result from processed images, a modified image will yield different 3D-objects. This process can be considered as *branching* and its resulting objects can be stored in a different segmentation node. Examples for branches could be a manual Ground Truth segmentation or automatic segmentations with different

```
<subject_database>
  <meta_data>... </meta_data>
  <data_directory> ... <data_directory>
  <data region="inguinal">
    <images>
      <raw id="" type="Angiography"> ... </raw>
      <processed id=""> ... </processed>
    </images>
    <objects>
      <segmentation id=""> ... </segmentation>
      <segmentation id=""> ... </segmentation>
    </objects>
  </data>
  <data region="spinal"> ... </data>
</subject_database>
```

Fig. 2. Basic XML database structure. All data definition nodes use an `id` attribute that is referenced in the data dictionary.

```
<data_directory>
  <structure type="">
    <structure type="">
      <element id="" mode="" name="" description=""> ... </element>
    </structure>
    <element id="" mode="" name="" description=""> ... </element>
  </structure>
</data_directory>
```

Fig. 3. Generic data directory structure. `type` and `name` attributes of structures and elements are application-specific. The `id` attribute creates a bijective mapping to the actual data object definition later in the document. The `mode` attribute specifies the available representations of a data object, e.g., for visualization and simulation. The `description` attribute stores a string that can be presented to the user, while the `name` attribute is used internally as a part of the file name of that object.

algorithms. A mere correction or update of an image also results in modified 3D-objects, which can now be stored with a new version number (Fig. 2). In order to organize changes of data, the authors propose a versioning of the XML file with a tool like Subversion. Thus, a technical user can benefit from the human readable format of the file while it does not grow with changes and still allows unrestricted access to its history.

3 Results

As a proof of concept, the approach was applied in a project for simulation of regional anaesthesia. Datasets consist of imaging data (DICOM series), 3D polygonal data (Wavefront OBJ files), other data (e.g., spline-based nerve cords hierarchy) and meta-information files (XML). The database is accessed by segmentation tools [6] and a VR-based simulator [7]. As a standardized naming convention for anatomical objects in the data directory the authors use entries from the FMA ontology.

Currently, the database holds five patient datasets, each with two original MR images (morphology, angiography) and about 50 anatomical structures (bones, muscles, etc.) as segmented image data and extracted 3D-objects.

The authors implemented a C/C++ library which abstracts the access patterns to the database and can be used with our segmentation, simulation and visualization tools to access data from within the application.

4 Discussion

A novel approach for medical data organization, called MEDOX, was described. The data objects are stored as files in the file system and referenced via an XML database file. This file contains a data directory that can be used to describe the application specific data dependencies by an application to interactively determine the available data objects without parsing the complete database. The XML file can easily be processed by established tools and APIs and integrated

```
<segmentation id="">
  <object id="" ref_id="" version="">
    <filename type="obj"> file_name_version.ext </filename>
    <properties> ... </properties>
  </object>
  <object id="" ref_id="" version=""> ... </object>
</segmentation>
```

Fig. 4. Structure of a segmentation node. The `ids` of the contained objects correspond with the `ids` in the data directory. `ref_id` corresponds with the `id` attribute of the image node the 3D-object is extracted from. The `version` number of an object is added as a suffix to the object's filename.

into own processing tools. As XML is an open standard, the user does not depend on proprietary third party tools. He can also add own extensions to the XML structure without rendering existing implementations useless.

The versioning of data objects is handled in two ways. The XML database that references the data files is versioned via SVN. The data files itself are not versioned. Storing the version history of a data file in the XML database would unnecessarily increase the database size and slow down its parsing. Instead, the versioned files are saved as copies with the according version number as a suffix in its filenames. Every version of the XML database references the most recent data file versions at that time.

Further development can concentrate on synchronisation of databases. The global database can be exported into a local copy (for performance reasons) that contain only a subset of the data objects. The local XML database is adjusted accordingly to reference only the available files. It is planned to support the import of new data from a local copy into the global database.

Additionally, new emerging standards like X3D and its MedX3D extension are observed and will be adapted into the approach once their drafts are finalized and getting certified.

Acknowledgement. This work was developed under the auspices of the German Research Foundation (DFG, RO 2000/7-1, KU 1132/4-1, LE 1108/8-1).

References

1. ISO. ISO/IEC 9075-1:2008: Information technology – Database languages – SQL – Part 1: Framework (SQL/Framework). Nederlands Normalisatie Instituut: pub-ISO; 2008.
2. Daly L, Brutzman D. X3D: extensible 3D graphics standard. In: Proc SIGGRAPH. New York, NY, USA: ACM; 2008. p. 1–6.
3. John NW, Aratow M, Couch J, et al. MedX3D: standards enabled desktop medical 3D. *Stud Health Technol Inform.* 2008;132:189–94.
4. Rosse C, Mejino JLVJ. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. *J Biomed Inform.* 2003;36(6):478–500.
5. Subversion. Subversion; 2006. <http://subversion.tigris.org>.
6. Teich C, Liao W, Ullrich S, et al. MITK-based segmentation of co-registered MRI for subject-related regional anaesthesia simulation. *Proc SPIE.* 2008;6918:2M–1–10.
7. Grottke O, Ntoubas A, Ullrich S, et al. Virtual reality-based simulator for training in regional anaesthesia. *Br J Anaesth.* 2009;103(4):594–600.

A Design Toolbox to Generate Complex Phantoms for the Evaluation of Medical Image Processing Algorithms

Omar Hamo, Georg Nelles, Gudrun Wagenknecht

Central Institute for Electronics, Research Center Juelich, D-52425 Juelich, Germany
g.wagenknecht@fz-juelich.de

Abstract. In the field of medical image processing, the evaluation of new algorithms is often a difficult task since real data sets do not allow a quantitative evaluation of the algorithms' properties and the correctness of results. Thus, a phantom design toolbox was developed to enable the generation of complex geometries appropriate to simulate anatomical structures as well as realistic image intensity properties and artifacts, such as noise and inhomogeneities. This paper describes the most important features of the new toolbox and shows sample phantoms generated so far.

1 Introduction

Generating appropriate simulated image data for the evaluation of medical image processing algorithms, such as segmentation [1] or registration, is often a challenging and time-consuming task, particularly in case of complex geometries of anatomical regions like the brain. So far, some evaluation databases [2, 3] exist which are based on one or more real image data set for which the expert segmentation result is provided and can be used as a gold standard. In addition, the BrainWeb [4] data set can be modified in image intensity values and image artifacts to build realistic brain phantoms whereas the geometry is fixed. To evaluate or test the behaviour of algorithms under different conditions, phantoms are needed which can be adapted in geometry as well as image intensities and artifacts to be suitable to given evaluation tasks. The phantom design toolbox was developed to simplify the creation of such custom-made and complex mathematical phantoms [5].

2 System description

The toolbox is provided with a graphical user interface based on Qt 4.5 [6] to easily choose, transform and combine geometrical primitives based on constructive solid geometry (CSG) [7]. Three orthogonal 2D views and one 3D view enable the user to monitor the phantom building process step-by-step.

2.1 Available primitives

So far, various rotationally symmetric and polygonal 2D and 3D primitives, as shown in Fig. 1, are available. In addition to these simple geometric primitives, a primitive called three dimensional supershape is made available. This primitive is based on the “Superformula” introduced by Gielis [8] designed “to study natural forms and phenomena”. Besides known approaches to create 3D supershapes by extrusion and the spherical product [9], a new approach is made available (1), (Fig. 2).

$$r(\phi, \theta) = \left[\left| \frac{\cos(m\phi/4)}{a} \right|^{n_2} + \left| \frac{\sin(m\phi/4)}{b} \right|^{n_3} + \left| \frac{\cos(m\phi/4)}{c} \right|^{n_4} \right]^{\frac{-1}{n_1}} \quad (1)$$

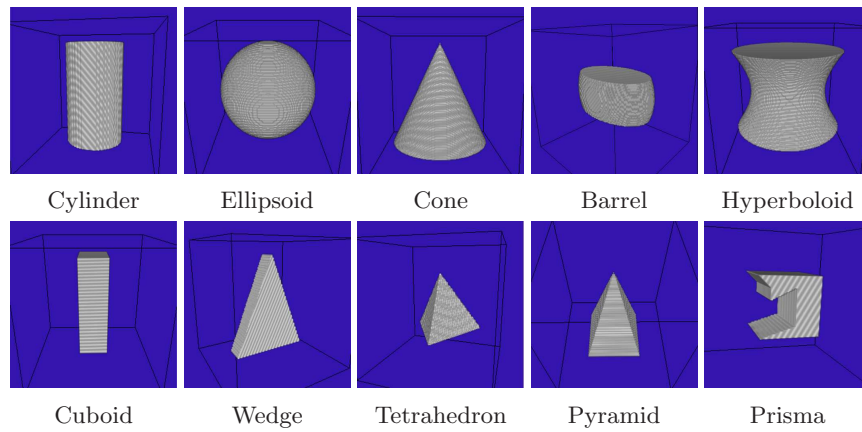


Fig. 1. 3D primitives.

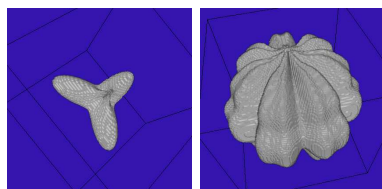


Fig. 2. Three dimensional supershape. By variation of only a few parameters, a variety of primitives can be created.

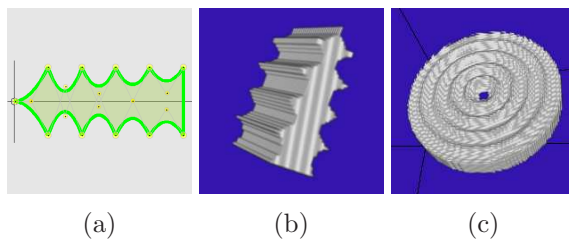


Fig. 3. Extrusion of a shape.

In order to generate even more complex and freely definable primitives, the toolbox provides a tool to draw two-dimensional boundary contours based on linear, quadratic and Bezier curves. The voxelized two-dimensional object (Fig. 3(a)) can then be extruded in one direction (Fig. 3(b)) or around one axis (Fig. 3(c)).

2.2 Composition of phantoms

The primitives chosen can be defined in value, size and position and be transformed by affine transformations. The boundary representation of the primitives is used to transform and position them on the volume grid. Phantoms can be composed of primitives using the Boolean operations union, difference, intersect or distinguish. The intensity value of the resulting combination can be calculated from the input values by operations, such as sum, difference or product. The resulting combination can in turn be used like a single primitive in further steps and be integrated into a user specific library.

In addition, the discrete values of the components can be modified to simulate real images including definable image artifacts, such as additive white Gaussian noise, and different kinds of inhomogeneities, such as nonlinear intensity non-uniformity fields from estimated real MRI scans [4]. In order to separate a phantom into different tissue types (e.g., white and grey matter for brain phantoms) dependent on the distance to the phantom's surface, an operation based on a distance transformation is made available.

The phantoms can be saved in all possible big and little endian data formats and different file formats, e.g., RAW and Analyze 7.5. The phantom exports module allows further histogram-based modifications of already assigned values, e.g., clipping or histogram equalization. Furthermore, the whole phantom building process can be saved in an editable XML file, to enable the user to stop and rebuild or modify the phantom at each step of the building procedure.

3 Results

The graphical user interface of the resulting phantom design toolbox is shown in Fig. 4. The phantom shows different brain tissue types and the simplified geometry of a branching cortical sulcus.

In Fig. 5, a phantom representing the corpus callosum is shown. It was created by using a MRI image of the sagittal view as a template to draw the border strip with cubic and Bezier curves. The voxelized and extruded primitive has then been combined with several hyperbolic, cylindrical and elliptical primitives using the CSG operation difference (A/B) to generate the curved surface shape.

Fig. 6 shows a phantom which is generated by drawing the contours of the cerebral cortex within the blue box. The gray matter was simulated by thickening and labeling these contours with a darker gray value. Then the voxelized object has been extruded by rotating it. The liquor is represented by an ellipsoid which was combined with the extruded primitive by the CSG operation union. A lesion

(green arrow) was simulated by a local spherical gradient field. Finally, noise and a nonlinear inhomogeneity field were applied.

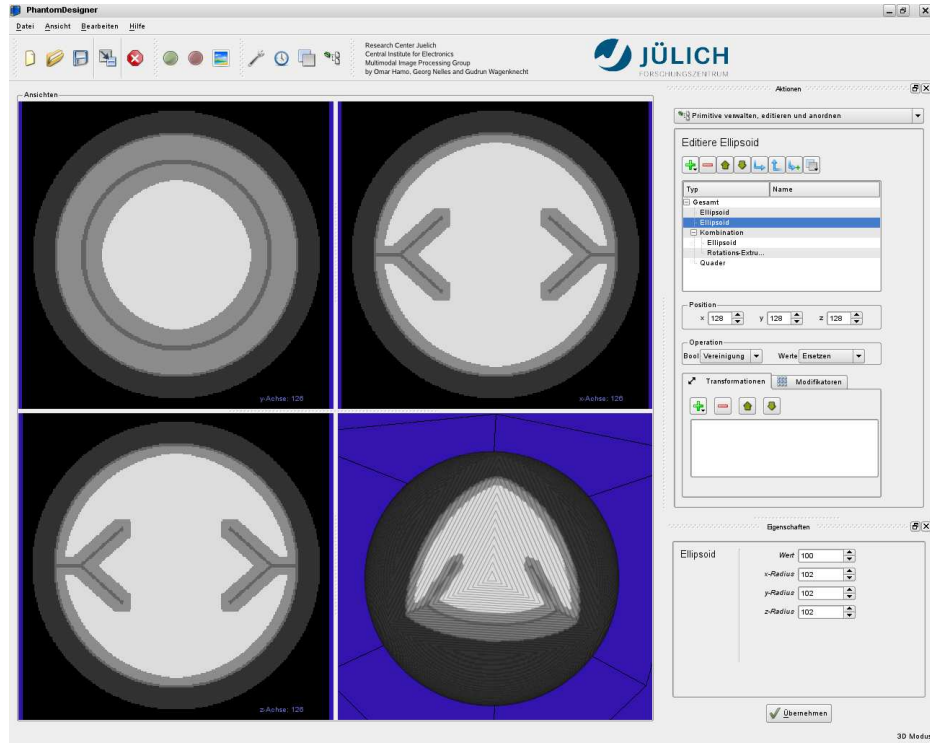


Fig. 4. Graphical user interface .

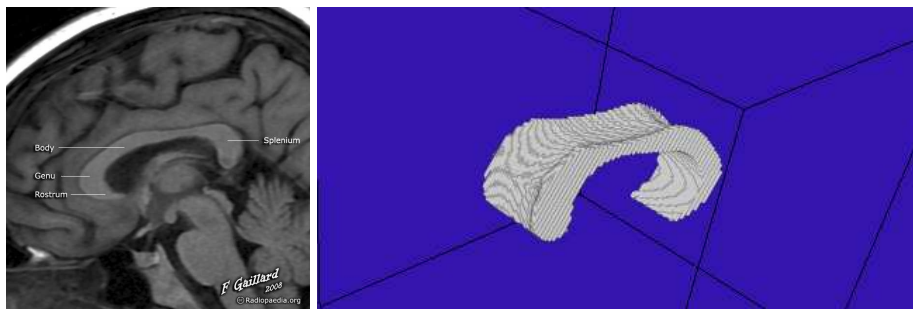
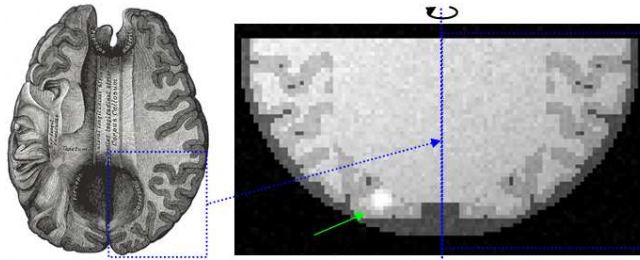


Fig. 5. Phantom (right) derived from a sample slice of the corpus callosum [10] (left).

Fig. 6. The phantom (right) is derived from a sample slice of the brain [11] (left).



4 Conclusion

Simulating complex geometries of anatomical regions is a difficult and time-consuming task. The phantom design toolbox allows building such complex phantoms in a fast, intuitive and easy way. Pairs of realistic-looking phantoms, one representing the anatomical structures and one representing the corresponding gray value image (e.g., of a simulated MR dataset), can be created to evaluate medical image processing algorithms. A second toolbox is already implemented to enable those quantitative evaluations.

Acknowledgement. This work was supported in part by the EC under Grant No. LSHC-2004-503569 and the BMBF under Grant No. 01EZ0822.

References

1. Wagenknecht G, Winter S. Volume-of-interest segmentation of cortical regions for multimodal brain analysis. Conf Record IEEE NSS/MIC. 2008; p. M06–455.
2. Kennedy DN, Worth AJ, Caviness VS. MRI-Based internet brain segmentation repository. Proc ISMRM. 1996;4(3):1657.
3. Rex DA, Ma JQ, Toga AW. The LONI pipeline processing environment. Neuroimage. 2003;19:1033–48.
4. Cocosco CA, et al. BrainWeb: Online interface to a 3D MRI simulated brain database. NeuroImage. 1997;5(4):425.
5. Hamo O, Nelles G, Wagenknecht G. A phantom design toolbox to generate simulated data suitable for the evaluation of segmentation algorithms. Proc IFMBE. 2009;25:PD 64.
6. Qt 4.5: Qt's Classes. Nokia Corporation and/or its subsidiaries; 2009 [cited 2009 Aug 16]. Available from: URL: <http://doc.trolltec.com/4.5/classes>.
7. Watt A. 3D-Computergrafik. vol. 3. Pearson Studium; 2003.
8. Gielis J. A generic geometric transformation that unifies a wide range of natural and abstract shapes. American J of Botany. 2003;90(3):333–8.
9. Bourke P. Supershapes in 3D [Online]; 2003 [cited 2009 Okt 16]. Available from: URL:<http://local.wasp.uwa.edu.au/~pbourke/geometry/supershapes3d/>.
10. Arcadian. Gray733. Wikipedia; 2007 [cited 2009 Okt 10]. Available from: URL: <http://commons.wikimedia.org/wiki/File:Gray733.png>.
11. Gaillard. Corpus-callosum. Wikipedia; 2008 [cited 2009 Okt 10]. Available from: URL: <http://upload.wikimedia.org/wikipedia/commons/1/12/Corpus-callosum.png>.

Multiscale Blood Vessel Segmentation in Retinal Fundus Images

Attila Budai¹, Georg Michelson², Joachim Hornegger¹

¹Pattern Recognition Lab and Graduate School in Advanced Optical Technologies(SAOT), University of Erlangen-Nürnberg, Erlangen, Germany

²Department of Ophthalmology, University of Erlangen-Nürnberg, Erlangen, Germany

attila.budai@informatik.uni-erlangen.de

Abstract. Retinal fundus imaging is widely used for eye examinations. The acquired images provide a unique view on the eye vasculature. The analysis of the vasculature has a high importance especially for detecting cardiovascular diseases. We present a multiscale algorithm for automatic retinal blood vessel segmentation, which is considered as a requirement for the diagnosis of vascular diseases. The algorithm uses a Gaussian resolution hierarchy to decrease computational needs, and allows to use of the same methods to detect vessels of different diameters. The algorithm is tested on two public databases using a common notebook. The algorithm segmented each image in less than 20s with a competitive accuracy over 93 % in both cases. This proves the applicability for medical applications.

1 Introduction

The images of the eye background provide the unique opportunity to in vivo examine the human vasculature. The sight of the vessel structure facilitates the detection of substantial vascular abnormalities e.g. aneurysms, artery/vein ratio [1]. Most of the automatic and semi-automatic methods to analyse the vessel structure rely on a vessel segmentation.

Segmentation of blood vessels is a research area since years, the current algorithms usually use some kind of vessel enhancement or feature extraction before the thresholding, or vessel tracking [2]. The methods with the highest accuracy also have high computational needs if thick vessels are present. The use of the proposed resolution hierarchy makes it possible to detect these vessels faster, while preserving a high accuracy.

We present a fast and accurate multiscale vessel segmentation algorithm to segment the vessels for further analysis. The segmentation algorithm is evaluated using two public databases [3, 4].

2 Materials and Methods

In general, our method works as follows:

- A Gaussian pyramid is generated.
- An efficient neighbourhood analysis applied on each level of the pyramid to enhance vessels of different diameter.
- The results of different levels are binarized and fused to obtain a final segmentation.

2.1 Gaussian Resolution Hierarchy

A Gaussian resolution hierarchy (Gaussian pyramid) is generated from the green channel of the fundus image, as it provides a reasonable contrast between the vessels and the background. The hierarchy consists of three levels (0–2). The original image has the highest resolution (level 0), and each further level has a halved width and height.

Due to the downscaling this hierarchy increases the speed of the segmentation, and allows the usage of the same 3×3 neighbourhood analysis to extract vessels with different diameters (Fig. 1). The neighbourhood analysis is applied on each level of the pyramid. This provides three grayscale images, whose pixels encode the probability, that the given pixel belongs to a vessel.

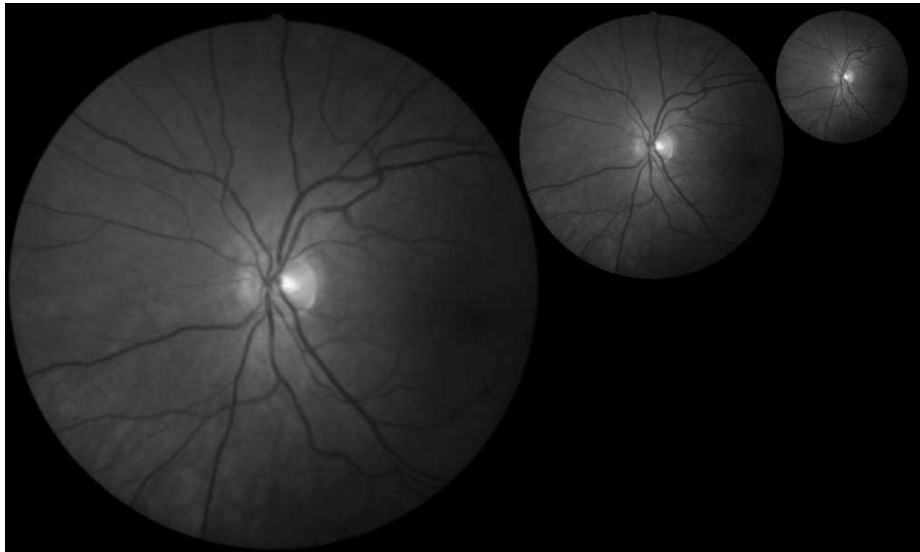


Fig. 1. The Gaussian resolution hierarchy generated from the green channel by decreasing resolution.

2.2 Neighbourhood Analysis

The Hessian matrix of the 3×3 neighbourhood for each pixel of the image are calculated. Since our images are two-dimensional, The hessian matrix will be a 2×2 matrix containing the second order derivatives

$$H(f) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} \quad (1)$$

The two eigenvalues of the matrix are calculated. They are reflecting the scale of the lowest curvature and the highest one in the neighbourhood. The value of the center pixel in the output layer is calculated from the ratio of these two eigenvalues

$$P_{\text{vessel}} = 1 - a_l/a_h \quad (2)$$

Where P_{vessel} is the result value for the given pixel, a_l is the lower, and a_h is the higher eigenvalue. In case of a vessel pixel the two eigenvalues representing the

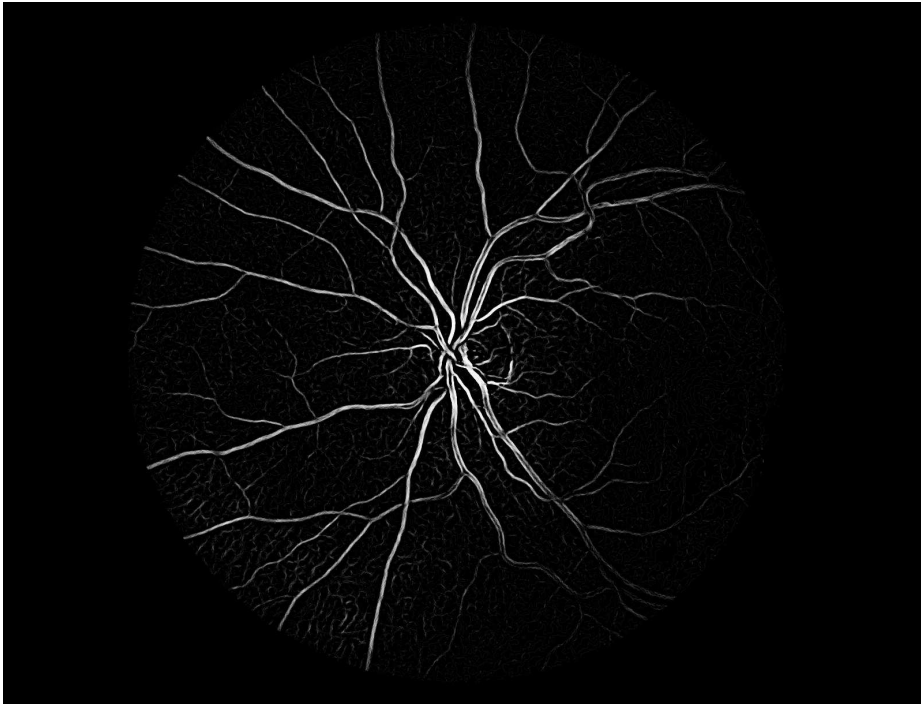


Fig. 2. A vessel enhanced image generated from the highest resolution level of the Gaussian hierarchy. The result was computed using the introduced neighbourhood analysis.

curvature are highly diverse and result in a ratio unlike one. In homogeneous regions the two eigenvalues are similar and result values close to one. This will cause a higher pixel value in the enhanced images if a vessel is present compared to the homogeneous regions [5]. This analysis is applied in each level of the pyramid. In Fig. 2 we show one of the three images of the enhanced vessels.

2.3 Image Fusion

All of the vessel-extracted images are now upscaled to the original image size to have a standard resolution. All images are binarized applying a hysteresis threshold, where the two thresholds are set manually by an expert to 85 % and 93 %. These two values showed the best visual results. The final vessel segmented image is achieved by merging. The binary images are merged by a pixel-wise OR operator (Fig. 3).

3 Results

The algorithm is evaluated using two public databases with gold standard manual segmentations. Both databases contain 20 retinal fundus images. The

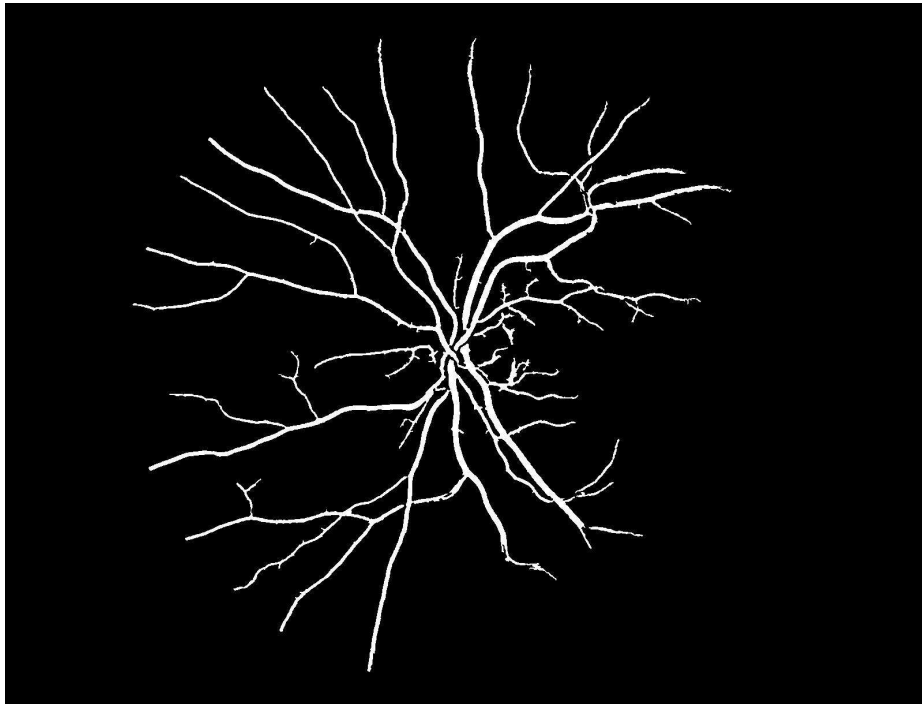


Fig. 3. The finally segmented image generated by binarizing and fusing the three vessel enhanced images.

Table 1. Evaluation using the two public databases.

Algorithm	Accuracy	Specificity	Sensitivity	Calculation time
Proposed method(STARE)	93.8%	97.5%	65.1%	16 sec
Proposed method(DRIVE)	94.9%	96.8%	75.9%	11 sec
Best published(DRIVE)	94.4%	-	-	-
Observer(DRIVE)	94.7%	-	-	-

STARE database[3] contains only the input and gold standard images with a resolution of 700×605 . The DRIVE database [4] contains 565×584 resolution images and additionally provides manual segmentations, which are done by a trained independent human observer and accuracy measurements of published algorithms. The test hardware was a notebook with a 2.0 GHz processor and 2.0 GB SDRAM. Tab. 1 shows the comparison of the proposed method to the human observer and the best published method of the DRIVE database.

4 Discussion

Based on the images of the mentioned public databases, we achieved a competitive accuracy compared to the other segmentation techniques or a human observer, while the applied resolution hierarchy decreases the computational needs. Our evaluations confirmed that the method offers a fast and reliable blood vessel segmentation applicable for quantitative analysis of vascular diseases.

References

1. Michelson G, Groh M, Groh M, et al. Telemedical screening of retinal vessels ("Talking Eyes"). Proc 102 DOG (Deutsche Ophthalmologische Gesellschaft). 2004.
2. Kirbas C, Quek F. A review of vessel extraction techniques and algorithms. ACM Comput Surv. 2004;36(2):81–121.
3. McCormick B, Goldbaum M. STARE = Structured Analysis of the Retina: Image processing of TV fundus image. Jet Propulsion Laboratory, Pasadena, CA: USA-Japan Workshop on Image Processing; 1975.
4. Staal JJ, Abramoff MD, Niemeijer M, et al. Ridge based vessel segmentation in color images of the retina. IEEE Trans Med Imaging. 2004;23(4):501–9.
5. Frangi AF, Frangi RF, Niessen WJ, et al. Multiscale vessel enhancement filtering. Lect Note Comp Sci. 1998;1496:165–72.

Scale-adaptive Wavelet-based Particle Detection in Microscopy Images

Oliver Greß¹, Birgit Möller¹, Nadine Stöhr², Stefan Hüttelmaier²,
Stefan Posch¹

¹Institute of Computer Science, Martin Luther University Halle-Wittenberg

²ZAMED, Martin Luther University Halle-Wittenberg

`oliver.gress@informatik.uni-halle.de`

Abstract. Stress granules and processing bodies play a major role in analysing the physiology of cells under various environmental conditions. We present a fully automatic approach to detect such particles in fluorescence labeled microscope images. The detection is based on scale-adaptive analysis of wavelet coefficients allowing for an accurate detection of particles with a large variety in size. Results on real images illustrate the appropriateness of our approach and proof high quality.

1 Introduction

Systems biology on the cellular level requires detailed analysis of different particles in cells like stress granules (SGs) and processing bodies (PBs). They are suggested to be dynamically linked to places of mRNA sorting and storage or degradation [1]. To understand and clarify the physiological roles of SGs and PBs, it is important to investigate the alterations of number, size, shape or contacts of these particles under different physiological conditions. Consequently the fully automatic detection of SGs and PBs is an essential tool to gain deeper insights into their biological role and function.

To detect spot-like particles, global and local thresholding techniques are still used in microscopy [2, 3]. Further techniques include h-dome transform followed by clustering [4]. Level set methods [5] on the other hand assume an approximately constant grey level shared by all target entities which typically does not hold for microscopy images. In [6, 7] a method to detect particles in microscopy images based on wavelet coefficients is proposed, but best-suited to detect particles with limited variation in size. Wavelet-based approaches are assumed to be superior to Fourier-based ones as their basis functions have local support.

We present an extension to the approach in [6, 7] aiming at the variable size of target entities by automatic scale-adaptation and test-based hypothesis selection.

2 Materials and Methods

The method in [6] is based on multi-scale analysis of wavelet coefficients. The original image $I_0(x, y)$ is recursively smoothed yielding images $I_1(x, y), \dots$,

$I_S(x, y)$. Wavelet coefficients $W_s(x, y)$ are derived as

$$W_s(x, y) = I_s(x, y) - I_{s-1}(x, y), \quad s \in \{1, \dots, S\} \tag{1}$$

For denoising, the amplitude-scale-invariant Bayes estimator [8] is applied, yielding adjusted coefficients $\tilde{W}_s(x, y)$. Features are represented by wavelet coefficients of adjacent scales. As the coefficients are correlated across scales due to the nature of the wavelet transform applied, adjacent scales are combined to a correlation image

$$c_{[a,b]}(x, y) = \prod_{s=a}^b \tilde{W}_s(x, y) \tag{2}$$

This correlation image is globally thresholded and the resulting connected components yield the final particles detected. The interval of scales $[a, b]$ used to correlate the wavelet coefficients defines the scales at which the particles of interest are represented. If all particles share the same characteristics, one interval is appropriate. In other cases one single interval either includes irrelevant scales or excludes important ones, in both cases often yielding incorrect particle size or shape, or missing particles.

To overcome these shortcomings we propose a new scale-adaptive technique which applies the wavelet-based segmentation to a set of – usually overlapping – intervals $[a_n, b_n]$, corresponding to different scales of particles. If for a target particle an appropriate interval is available, it is usually correctly detected in this interval. However, in many cases the particle is also found in adjacent intervals with incorrect size or shape (Fig. 1). This results in the necessity to select the correct one from overlapping and, thus, competing particle hypotheses. For this we propose an approach based on statistical hypothesis testing.

Typically hypotheses from the correlation image of a coarse scale result in larger regions than hypotheses from finer scales. Additionally, regions can be split up in finer scales due to the presence of multiple smaller particles or varying gray values inside larger particles.

We assemble competing hypotheses in trees whose nodes correspond to the particle regions detected in different intervals. The trees are build bottom-up

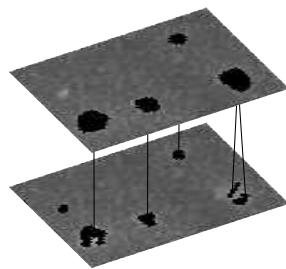


Fig. 1. Segmentation for two adjacent intervals (top: coarse, bottom: fine) and resulting hypotheses trees.

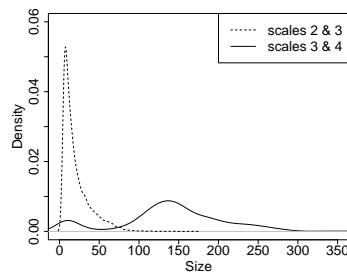


Fig. 2. Distribution of granule size detected in intervals $[a_1, b_1]$ and $[a_2, b_2]$.

starting with particle hypotheses from the finest scale. At each coarser level edges are inserted between a region and all overlapping regions of the next finer scale. Fig. 1 gives an example of the resulting hypotheses trees for two intervals.

In principle, two different regions at a coarse scale may both overlap with the same region at the next finer scale, destroying the tree characteristics of the hypothesis graph. However, this happened only once for the whole test data. Such constellations can for example be resolved using a simple criterion of largest overlap. The selection of one hypothesis out of several competing ones is accomplished using again a bottom-up procedure, starting from the fine scale leaves. Parent nodes are compared with their children and inferior nodes are deleted.

For comparison we employ the concept of meaningful events ([9]). This concept is tightly related to statistical hypothesis testing. In our case the null hypothesis H_0 models the case where no real particle is present at the location to be analyzed, rather a particle was detected due to noise or chance. To compute the likelihood $P(F_i | H_0)$ of a particle F_i detected under H_0 , the observations at all pixels are assumed to be pairwise independent

$$P(F_i | H_0) = \prod_{(x,y) \in F_i} P(C_{[a_n, b_n]}(x, y) = c_{[a_n, b_n]}(x, y) | H_0) \quad (3)$$

where $C_{[a_n, b_n]}(x, y)$ are random variables modelling the correlation value observed at position (x, y) . $P(C_{[a_n, b_n]}(x, y) = c_{[a_n, b_n]}(x, y) | H_0)$ is the probability to observe the value $c_{[a_n, b_n]}$ at location (x, y) due to noise. Following [10] we estimate $P(C_{[a_n, b_n]}(x, y) | H_0)$ as the discretized histogram of the complete correlation image for interval $[a_n, b_n]$.

The p-value $p(F_i)$ of F_i is the probability to observe a particle under the null hypothesis with correlation values at least as extreme as the ones of F_i . I.e., a particle with extremier values has at each pixel location in the correlation image a value larger than the one observed for F_i .

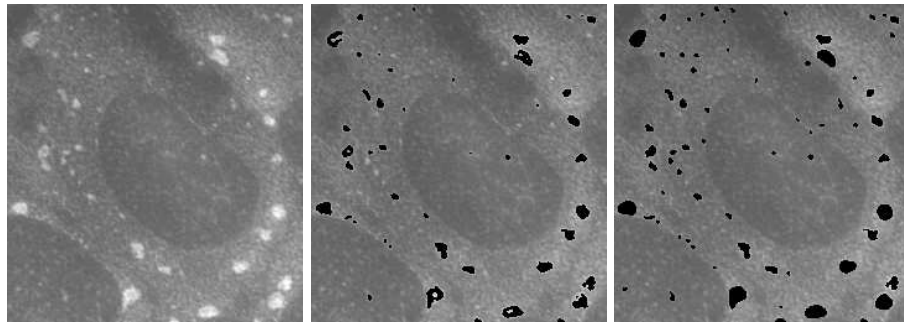


Fig. 3. Detection of stress granules. Left: Detail of image no. 8 with stress granules of varying sizes. Middle: Detection results using the method [6]. Right: Detection results using the scale-adaptive method (Image values scaled for visualization of results).

Still assuming independence of pixels this yields

$$p(F_i) = \prod_{(x,y) \in F_i} P (C_{[a_n, b_n]}(x, y) \geq c_{[a_n, b_n]}(x, y) \mid H_0) \quad (4)$$

We use this concept to compare a set of overlapping particle hypotheses to delete inferior nodes from the trees. We keep the particles with smallest p-value which consequently are assumed to be the ones most unlikely caused by chance. As p-values of particles with different size of support are compared, these raw p-values are normalized according to their support to allow fair comparison. In the case of multiple children their p-values are multiplied for comparison with the parent and we decide for F_i at the coarser scale if

$$p(F_i)^{\frac{1}{|F_i|}} < \prod_{\{k \mid F_k \text{ child of } F_i\}} p(F_k)^{\frac{1}{|F_k|}} \quad (5)$$

and for particles F_k on the finer scale otherwise.

3 Results

The proposed approach is tested on 10 microscope images of U2OS osteosarcoma cells stressed with sodium arsenate for one hour before fixation. SGs were labeled by immunostaining of ZBP1 in red, and for 5 images PBs were labeled by immunostaining of DCP1A in green.

For our application, two overlapping intervals of scale $[a_1, b_1] = [2, 3]$ and $[a_2, b_2] = [3, 4]$ have shown to be sufficient and will be used for the experiments reported in the following. Fig. 3 shows a detail of image no. 8 with fluorescently labeled SGs, and segmentation results for the method [6] with scales $[a, b] = [2, 4]$ and the proposed scale-adaptive method, respectively. For the scale-adaptive method the distribution of the size of granules detected in each of the scale intervals is depicted in Fig. 2. In analogy to Fig. 3, Fig. 4 gives segmentation results for PBs in a part of image no. 2.

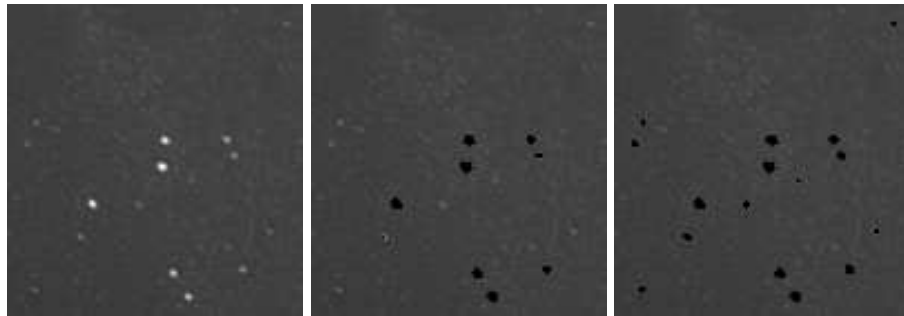


Fig. 4. Detection of processing bodies. Left: Detail of image no. 2 with processing bodies. Middle: Detection results using the method [6]. Right: Detection results using the scale-adaptive method. (Image values scaled for visualization of results).

4 Discussion

Stress granules show a large variety of different sizes and shapes. Using only one interval of wavelet coefficients as in [6] imposes the implicit constraint on similar shape and size for all granules. As can be seen from Fig. 3 (middle) this allows to detect a set of pronounced granules, however, misses several smaller granules and sometimes leads to incomplete segmentation for large granules, i.e. their shapes show deep convexities unusual for granules. In contrast, applying our new scale-adaptive approach based on selection of detection results from different scale intervals overcomes this problem. The detection is improved as detected granules cover a larger range of different scales and have more accurate contours (Fig. 3, right). The local adaptivity of our method shows also in Fig. 2. Detections from interval $[a_2, b_2]$ correspond mainly to large-sized granules, while small granules are detected predominantly in interval $[a_1, b_1]$.

Detection results for PBs also demonstrate the ability of our approach to automatically select features from the best scale. Compared to the results of method [6] our detection also includes PBs of less saliency (Fig. 4) avoiding the canceling effect of coarse scales. The variance in size among PBs is smaller than among SGs. Accordingly, a single fine-scale interval should be sufficient to detect the majority of PBs. Indeed, although intervals $[a_1, b_1]$ and $[a_2, b_2]$ are used for PBs detection, 99 % of the PBs are selected from interval $[a_1, b_1]$.

References

1. Kedersha N, Stoecklin G, Ayodele M, et al. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *Cell Biol.* 2005;169(6):871–4.
2. Xavier J, Schnell A, Wuertz S, et al. Objective threshold selection procedure (OTS) for segmentation of scanning laser confocal microscope images. *J Microbiol Methods.* 2001;47(2):169.
3. Bolte S, Cordelieres FP. A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc.* 2006;224(3):213–32.
4. Smal I, Niessen W, Meijering E. A new detection scheme for multiple object tracking in fluorescence microscopy by joint probabilistic data association filtering. In: *Proc IEEE ISBI; 2008.* p. 264–7.
5. Chan TF, Vese LA. Active contours without edges. *IEEE Trans Image Process.* 2001;10(2):266–77.
6. Olivo-Marin JC. Extraction of spots in biological images using multiscale products. *Pattern Recognit.* 2002;35(9):1989–96.
7. Dufour A, Meas-Yedid V, Grassart A, et al. Automated quantification of cell endocytosis using active contours and wavelets. In: *Proc ICPR; 2008.*
8. Figueiredo MAT, Nowak RD. Wavelet-based image estimation: an empirical Bayes approach using Jeffrey’s noninformative prior. *IEEE Trans Image Process.* 2001;10(9):1322–31.
9. Desolneux A, Moisan L, Morel JM. A grouping principle and four applications. *IEEE Trans Pattern Anal Mach Intell.* 2003;25(4):508–13.
10. Desolneux A, Moisan L, Morel JM. Edge detection by helmholtz principle. *J Math Imaging Vis.* 2001;14(3):271–84.

A Cortex Segmentation Pipeline for Neurosurgical Intervention Planning

Daniela I. Wellein¹, Matthias Pfeifle², Martin Althuzes¹, Lars Voitel¹,
Dirk Bartz¹

¹VCM/ICCAS, Universität Leipzig

²Neurochirurgische Klinik, Universitätsklinikum Tübingen

daniela.wellein@iccas.de

Abstract. Cortex segmentation is a prerequisite for 3D neurosurgical intervention planning. The visualization of the cortical surface along with target and risk structures enables conservative access planning and gives context information about the patient-specific anatomy. We present a pipeline for the segmentation of the cortical surface in T1-weighted MR images. To segment the cortex, we combine watershed and level set segmentation. The cortex segmentation pipeline (CSP) is semiautomatic and designed to let the surgeon influence the intermediate data at any stage of the segmentation process. To evaluate the CSP we used the Segmentation Validation Engine (SVE) by Shattuck et al. [1] and compared the results to three different popular methods (BET, BSE and HWA).

1 Introduction

Along with risk and target structures, 3D neurosurgical intervention planning requires a visualization of the cortex. The sulci and gyri of the cortical surface provide anatomical landmarks which give the surgeon a better overview of the patient-specific anatomy. The course of the sulci and their relation to the target structure facilitates access planning along the sulci and therefore minimizes the destruction of healthy cortical tissue. Prerequisite for the visualization is a segmentation of the cortex. A variety of automated segmentation methods exist. Most popular are the brain surface extractor (BSE, [2]), brain extraction tool (BET, [3]) and the hybrid watershed algorithm (HWA, [4]). These methods let the user specify one to three parameters, perform all calculations, and show only the final result. In contrast, our multi-step pipeline allows the inspection (and correction) of all intermediate results. Therefore, the need for a time consuming trial and error process of finding the right parameters is reduced.

2 Materials and Methods

2.1 Pipeline

We integrated the cortical segmentation pipeline (CSP) into our volume data processing and visualization platform VolV [5]. The segmentation algorithms

are based on the Insight Segmentation and Registration Toolkit (www.itk.org). CSP consists of three main steps: Preprocessing, watershed, and threshold level set segmentation (Fig. 1). In the preprocessing step, we use anisotropic diffusion filtering [6] to reduce image noise while retaining edges. The next step utilizes the approach of Hahn et al. [7] to coarsely identify brain tissue in the volume datasets. Here, isovalues of the denoised dataset are inverted and serve as input for a watershed transform, leading to a set of basins. In most cases the cortex is included in one basin, which forms the initial brain mask. To close holes in the mask and reduce leaking, morphological operations are applied. Morphological closing is performed with a spherical kernel with radius $r = 2$ and erosion with a spherical kernel with radius $r = 4$.

In the last step, a refined brain mask is computed with threshold level set segmentation. Input data are given by the original volume dataset and the postprocessed initial brain mask from the watershed segmentation. The level set function deforms the initial brain mask according to the propagation term

$$P(x, y, z) = \begin{cases} g(x, y, z) - l, & \text{if } g(x, y, z) < (u - l)/2 + l \\ u - g(x, y, z), & \text{otherwise} \end{cases} \quad (1)$$

with $g(x, y, z)$ being the isovalue of an arbitrary voxel and l and u being the lower and upper threshold. A careful selection of the level set thresholds (2.2), leads to a brain mask without liquor and blood vessels. Postprocessing the brain mask, leaking is reduced using morphological opening (largest connected component).

2.2 Parameter specification

To let the user influence all intermediate results and simultaneously minimize unnecessary user interaction, we give reference values for the segmentation parameters. The first parameters are required for the selection of the watershed pre-flood height and the flood level. The pre-flooding parameter p serves to set all isovalues below $v = g_{\min} + p(g_{\max} - g_{\min})$ to v , with g_{\min} and g_{\max} being the minimal and maximal isovalue in the respective volume dataset. The flood level f controls the number of basins in the segmentation result. In the evaluated volume datasets, $p = 0.1$ and $f = 0.35$ lead to one basin enclosing the brain.

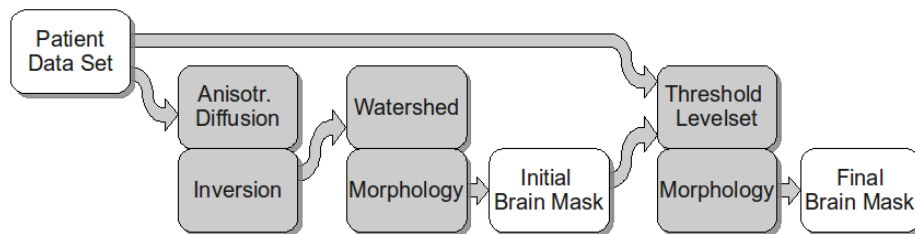


Fig. 1. Workflow of the segmentation pipeline.

Level set thresholds l and u are directly derived from the histogram of the volume dataset. For all cases, three specific points can be identified in the histogram: The local minimum (min), local maximum (max), and the steep slope (slp). In Fig. 2, three histograms with a manually fitted curve and the three points are being showcased. Level set thresholds are chosen, so that $l = \min + (\max - \min)/2$ and $u = \text{slp}$.

2.3 Evaluation Methodology

To evaluate the CSP, we applied it to 40 datasets provided in the SVE [1]. SVE compares the segmentation results (S) to a manual cortex segmentation (T) and computes four different quality measures: Sensitivity, Specificity, Jaccard similarity ($|S \cap T|/|S \cup T|$), and the Dice coefficient ($2|S \cap T|/(|T| + |S|)$). Additionally SVE provides error maps of average false positives and negatives and a table of the true positive rates for a variety of different cortical structures.

3 Results

Table 1 shows the results for the CSP compared to three of the most popular cortex segmentation methods. With default parameters, our pipeline achieved an accuracy in the range of the default parameters of BET, BSE and HWA. Mean values with standard deviation are: Jaccard = 0.9239 (± 0.014), Dice = 0.9604

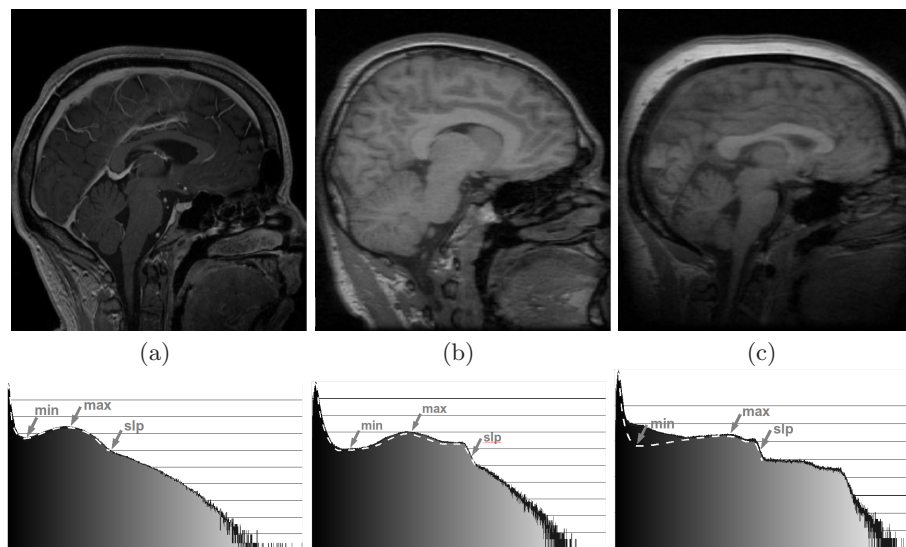


Fig. 2. Example histograms with the min, max and slp marked: (a) with contrast agent, (b) without contrast agent, (c) with signal inhomogeneities. The dashed line marks the course of manually fitted curves.

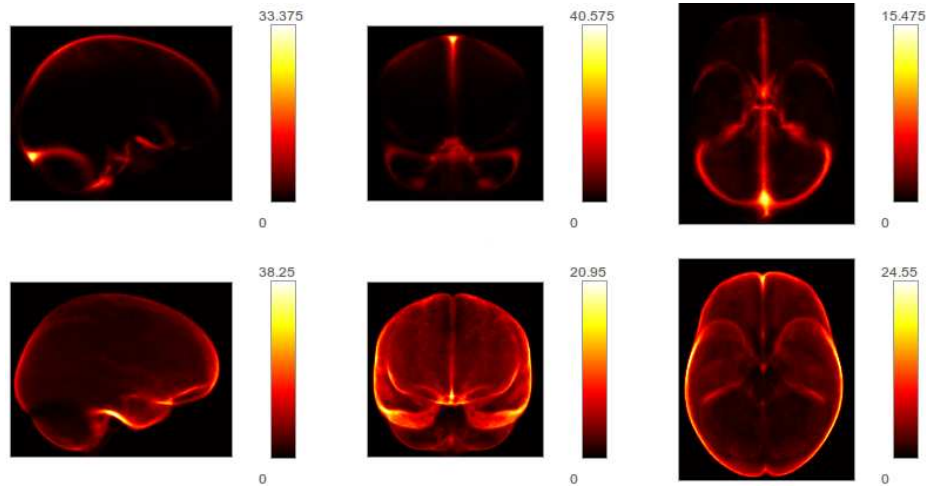
Table 1. Error metrics for CSP, BET, BSE, and HWA. Reference values from [1] are shown for the default parameters and the best parameters set run.

Method	Parameters	Jaccard	Dice	Sensitivity	Specificity
CSP	default	0.9239	0.9604	0.9543	0.9941
BET	default	0.8919	0.9420	0.9858	0.9804
	-B	0.9400	0.9691	0.9627	0.9957
BSE	default	0.5956	0.7272	0.9804	0.8538
	-n5 -d18 -s0.7 -p	0.9394	0.9684	0.9725	0.9937
HWA	default	0.8531	0.9207	0.9992	0.9693
	-less	0.8537	0.9210	0.9992	0.9695

(± 0.008), Sensitivity = 0.9543 (± 0.018), and Specificity = 0.9941 (± 0.003). The standard deviations indicate that CSP provides satisfactory results throughout all tested volume datasets.

4 Discussion

While Jaccard similarity and Dice coefficient fits in the range of the results of BET, BSE and HWA and specificity is pretty high, sensitivity of the CSP is comparatively low (Tab. 1). This especially occurs for datasets with signal inhomogeneities. As CSP is mainly isovalue based, we cannot deal with large signal

**Fig. 3.** Error maps for CSP show the average false positives (top) and negatives (bottom). False positives occur at the venous sinus and the eye sockets, false negatives mainly at the temporal lobe.

inhomogeneities. For those datasets and with no inhomogeneity correction our method has to trade leaking for undersegmentation - specificity for sensitivity.

A look at the error maps (Fig. 3) gives a more specific understanding where undersegmentation and leaking occurs. The separation of the venous sinus from the cortex was incomplete and in some cases the segmentation leaks into the eye sockets. Areas with brain tissue missing in the segmentation are located in the temporal lobe.

In the future we aim to reduce the problems of leaking and undersegmentation with an inhomogeneity correction and morphological operations. Also our threshold selection process could be automatized by fitting a curve to the histogram and deriving the *min*, *max*, and *slp* from the fitted curve.

References

1. Shattuck DW, Prasad G, Mirza M, et al. Online resource for validation of brain segmentation methods. *NeuroImage*. 2009;45(2):431–9.
2. Shattuck DW, Sandor-Leahy SR, Schaper KA, et al. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*. 2001;13(5):856–76.
3. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002;17(3):143–55.
4. Segonne F, Dale AM, Busa BE, et al. A hybrid approach to the skull stripping problem in MRI. *NeuroImage*. 2004;22:2004.
5. Pfeifle M, Born S, Fischer J, et al. VolV: Eine OpenSource-Plattform für die medizinische Visualisierung. In: *Proc CURAC*; 2007. p. 193–6.
6. Perona P, Malik J. Scale-Space and edge detection using anisotropic diffusion. *IEEE Trans Pattern Anal Mach Intell*. 1990;12(7):629–39.
7. Hahn HK, Peitgen HO. The skull stripping problem in MRI solved by a single 3D watershed transform. In: *Proc MICCAI*; 2000. p. 134–43.

Model-Based Lower Limb Segmentation using Weighted Multiple Candidates

André Gooßen¹, Eugen Hermann¹, Thorsten Gernoth¹
Thomas Pralow², Rolf-Rainer Grigat¹

¹Vision Systems, Hamburg University of Technology, 21079 Hamburg, Germany

²General X-Ray, Philips Healthcare, 22335 Hamburg, Germany

andre.goossen@tu-harburg.de

Abstract. In this paper we propose an extension of active shape model based bone segmentation. We examine the benefit of using multiple candidates for new landmark positions during segmentation. To incorporate this information we compare three strategies of adapting the fitting algorithm. For evaluation we segmented the hip, knee and ankle joints in more than 100 digital radiographs of the lower limbs. We achieve superior accuracy compared to the classic algorithm and prove that segmentation results benefit from multiple candidates.

1 Introduction

Segmentation of bone structures in digital radiographs is a prerequisite for quantitative orthopaedic examinations, such as length and angle [1], bone density or bone age measurements and preoperative planning [2]. However, an accurate delineation of these structures by experts is time-consuming [3] and often subject to intra- and inter-observer variability thus requiring a robust computer-aided segmentation capable of processing the large variety of images in clinical practice.

A well-known approach to segmenting digital radiographs are active shape models (ASM) [4]. However when it comes to modelling the appearance of medical image data it is often impossible to determine an accurate and robust feature. Spurious edges and noise produce additional candidates and might cause the segmentation to fail. Taking multiple candidates of a single or different image features [5] increases the probability of including the true position. We propose and evaluate three approaches for the task of segmenting the lower limbs in long-leg radiographs and compare the results in terms of accuracy and robustness.

2 Materials and Methods

In the original Active Shape Model (ASM) algorithm [4] the shape \mathbf{x} is modelled by adding a linear combination of eigenvectors $\mathbf{P} = (\mathbf{p}_1|\mathbf{p}_2|\dots|\mathbf{p}_t)$ describing possible shape variation to the mean shape $\bar{\mathbf{x}}$. The appearance is modeled using sampling vectors around each landmark $(x_i, y_i)^\top$, $i = 1, 2, \dots, n$. In an iterative manner new landmark positions (\hat{x}_i, \hat{y}_i) are estimated from the local appearance.

We call these positions candidates. A new shape \mathbf{x} is generated from these candidates by least-squares fitting, i.e. minimizing the error

$$\Delta = (\hat{\mathbf{x}} - (\bar{\mathbf{x}} + \mathbf{P}\mathbf{b}))^\top (\hat{\mathbf{x}} - (\bar{\mathbf{x}} + \mathbf{P}\mathbf{b})) \quad (1)$$

with \mathbf{b} denoting the parameters of the shape model.

However, this procedure results in outlying points disproportionately affecting the fit, which may or may not be desirable depending on the problem. The least-squares minimization yields optimal results for Gaussian distributed errors. Due to the non-uniform appearance and detailed structure in medical images the assumption of a Gaussian model for the residual distribution is seldom accurate. For that reason we add a quality measure for each candidate to control the candidate's influence on the fit. The quality of each candidate is represented in form of weights w_i , $0 \leq w_i \leq 1$ combined to a weight matrix

$$\mathbf{W} = \begin{bmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_{2n} \end{bmatrix} \quad (2)$$

with $w_{n+i} = w_i \forall 1 \leq i \leq n$. Using this extension, we derive the new error term

$$\Delta = (\hat{\mathbf{x}} - (\bar{\mathbf{x}} + \mathbf{P}\mathbf{b}))^\top \mathbf{W} (\hat{\mathbf{x}} - (\bar{\mathbf{x}} + \mathbf{P}\mathbf{b})) \quad (3)$$

This technique only considers single candidates per landmark. However, there might be additional information such as multiple local optima or candidates of different image features [5]. This follows the idea of boosting, i.e. taking many weak candidates and combining them into a robust and accurate result.

To incorporate this information we extend the fitting algorithm to multiple candidates per landmark and examine three possible strategies. Let κ_i denote the number of candidates for the i -th landmark and let $w_{i,k}$, $k = 1, 2, \dots, \kappa$ denote the corresponding weights:

- *Best/Fittest Candidate*: For each landmark we simply keep the candidate with the highest weight $w_i = \max(w_{i,k})$, regardless of the overall-rating of this weight compared to neighbouring landmarks. For an appropriate generation of the weights, this strategy always selects the fittest candidate. However, if all the weights of a certain landmark are small, this conflicts with the idea of boosting.
- *Multiple Candidates*: Instead of neglecting candidates with small weights, we integrate all candidates for a specific landmark into the error term. The new equations are given with respect to the candidate's coordinates. With κ_i denoting the number of estimations for the i -th landmark and the total number of feature points $\kappa = \sum_{i=1}^n \kappa_i$, the extended target vector $\hat{\mathbf{x}} \in \mathbb{R}^{2\kappa}$ is given by

$$\hat{\mathbf{x}} = (x_{1,1}, \dots, x_{1,\kappa_1}, \dots, x_{n,1}, \dots, x_{n,\kappa_n}, y_{1,1}, \dots, y_{1,\kappa_1}, \dots, y_{n,1}, \dots, y_{n,\kappa_n})^\top \quad (4)$$

The mean vector $\bar{\mathbf{x}} \in \mathbb{R}^{2\kappa}$ becomes

$$\bar{\mathbf{x}} = \left(\underbrace{x_1, \dots, x_1}_{\kappa_1}, \dots, \underbrace{x_n, \dots, x_n}_{\kappa_n}, \underbrace{y_1, \dots, y_1}_{\kappa_1}, \dots, \underbrace{y_n, \dots, y_n}_{\kappa_n} \right)^\top \quad (5)$$

In the same way each eigenvector \mathbf{p}_k , with $k = 1, \dots, t$ has to be extended to

$$\mathbf{p}_k = \left(\underbrace{p_1, \dots, p_1}_{\kappa_1}, \dots, \underbrace{p_n, \dots, p_n}_{\kappa_n}, \underbrace{p_{n+1}, \dots, p_{n+1}}_{\kappa_1}, \dots, \underbrace{p_{2n}, \dots, p_{2n}}_{\kappa_n} \right)^\top \quad (6)$$

- *RANSAC/MSAC*: This statistical approach [6] has the advantage of implicitly defining neither Gaussian nor any other error distribution and thus should work for any kind of candidate distribution, e.g. bimodal for spurious edges or equally distributed for image parts dominated by noise.

We derive a model by randomly sampling the minimum number of candidates t required to generate a shape. Subsequently we identify the number of candidates within a corridor τ around this shape. However we do not count, following the RANSAC algorithm, but rather sum up the weights of all candidates within this consensus set

$$C = \sum_i w'_{i,\kappa_i}, \quad w'_{i,\kappa_i} = \begin{cases} 0 & \delta((\hat{x}_{i,\kappa_i}, \hat{y}_{i,\kappa_i})^\top) > \tau \\ w_{i,\kappa_i} & \delta((\hat{x}_{i,\kappa_i}, \hat{y}_{i,\kappa_i})^\top) \leq \tau \end{cases} \quad (7)$$

Thus, the shape with the largest sum wins. This technique can be further extended; by selecting the candidates with a probability proportional to their weight, instead of equal probabilities, we enhance the chance to select a set of candidates on the true shape.

We applied the three strategies to the task of segmenting the joints in 109 clinical long-leg radiographs [7]. We measured the mean and maximum curve-to-curve (Fréchet) distance to manual expert delineations and compared the errors. To evaluate the lower bound for the accuracy we also quantified the inter-observer error and performed a study with five observers. The reference implementation is the classic ASM algorithm with an unweighted single candidate per landmark.

3 Results

The average processing time for a full segmentation (six joints) is 7 s for best candidate strategy, 11 s for multiple candidates and 30 s for RANSAC on a standard 4×3 GHz machine. Refer to Fig. 2 for a depiction of the overall quantitative results. The segmentation accuracy is increased for all three methods in terms of a lower mean as well as maximum error. For each of the partial segmentations of the whole leg, using multiple candidates results in smallest errors. The mean error is 0.59 mm for the hip, 0.47 mm for the knee, and 0.37 mm for the ankle joint compared to an inter-observer variability of 0.49 mm, 0.43 mm, and 0.35 mm respectively. Fig. 1 shows typical results of the three methods.

4 Discussion

We introduced and evaluated three approaches to incorporate multiple candidates for new landmark positions into classic ASM segmentation. Experiments have shown that weighted fitting of multiple candidates is superior to using single candidates or unweighted fitting.

It is evident from the mean error that the use of weighted multiple candidates considerably improves the segmentation accuracy for delineating bone structures in long-leg radiographs. The significant drop in the maximum error in all three methods indicates that, in addition to the higher accuracy, also the robustness is increased compared to classic ASM segmentation.

The remaining error is close to the inter-observer variability which forms the lower bound when measuring the accuracy. Moreover, further improvement of the segmentation is limited by the shape model's reconstruction error. It is thus

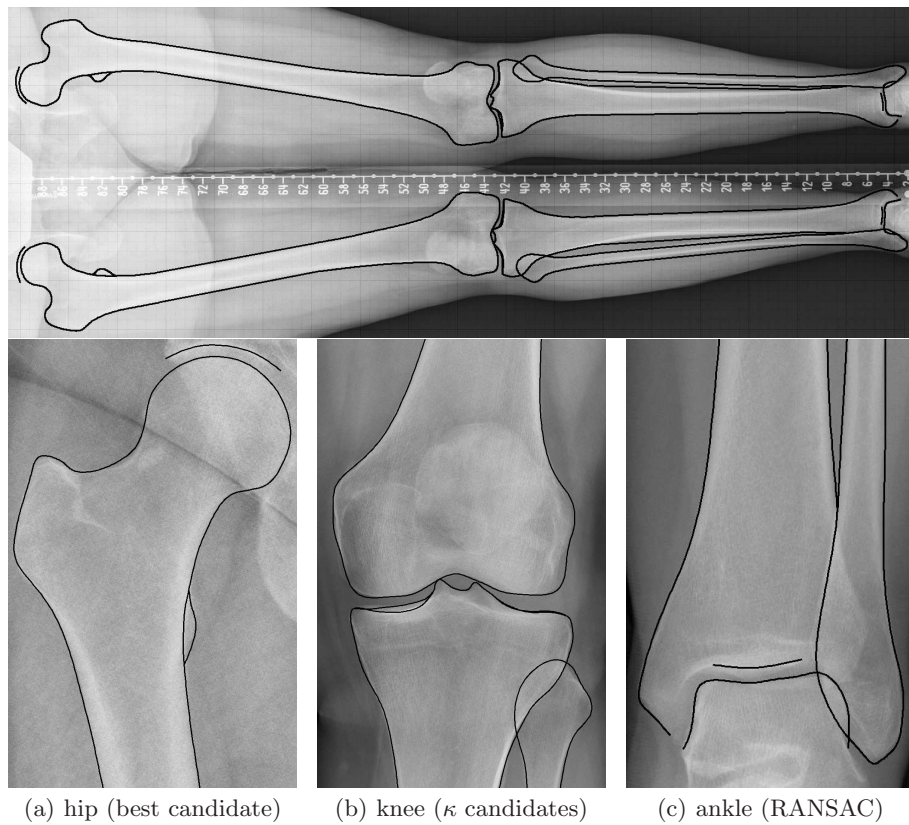
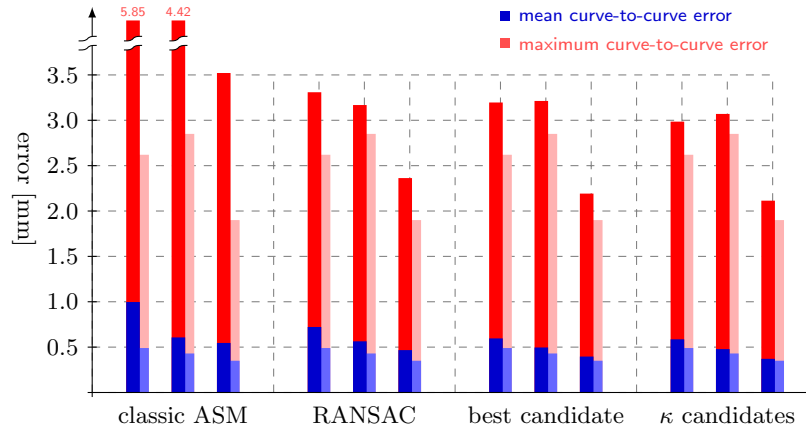


Fig. 1. Full leg segmentation in an oversized radiograph and typical delineation results of the segmentation for (a) hip joint (mean error 0.59 mm), (b) knee joint (mean error 0.47 mm) and (c) ankle joint (mean error 0.37 mm) using the three proposed methods.

Fig. 2. Mean and maximum curve-to-curve error for the classic ASM algorithm [4] and RANSAC strategy, weighted least squares with best candidate and multiple candidate strategy. The bars depict results for the hip (left), knee (mid) and ankle joint (right). The shaded bars in the background result from the inter-observer study and thus depict a measure for the accuracy of the manual delineation.



necessary to apply a separate post-processing to the ASM result, which will be addressed in future work.

References

1. Boewer M, Arndt H, Ostermann PW, et al. Length and angle measurements of the lower extremity in digital composite overview images. *Eur Radiol.* 2005;15(1):158–64.
2. Pafilas D, Nayagam S. The pelvic support osteotomy: indications and preoperative planning. *Strategies Trauma Limb Reconstr.* 2008;3(2):83–2.
3. Hochhausen C. Vergleich der software-unterstützten Vermessung von Röntgenaufnahmen der unteren Extremität im Rahmen der präoperativen Planung für Korrekturingriffe mit der manuellen Planung [PhD Thesis]. Medizinische Hochschule Hannover; 2004.
4. Cootes TF, Taylor CJ, Cooper DH, et al. Active shape models - their training and application. *Comput Vis Image Underst.* 1995;61(1):38–59.
5. Gooßen A, Peters D, Gernoth T, et al. Intelligent feature selection for model-based bone segmentation in digital radiographs. In: *Proc Int Technol Appl in Biomed.* IEEE; 2009. p. 1–4.
6. Rogers M, Graham J. Robust active shape model search. *Lect Notes Computer Sci.* 2002;2352:517–30.
7. Gooßen A, Schlüter M, Pralow T, et al. A stitching algorithm for automatic registration of digital radiographs. *Lect Notes Computer Sci.* 2008;5112:854–62.

Fully Automatic Model Creation for Object Localization utilizing the Generalized Hough Transform

Heike Ruppertshofen^{1,2,3}, Cristian Lorenz², Peter Beyerlein⁴, Zein Salah³,
Georg Rose³, Hauke Schramm¹

¹Institut für Angewandte Informatik, Fachhochschule Kiel

²Sector Medical Imaging Systems, Philips Research Europe - Hamburg

³Institut für Elektronik, Signalverarbeitung und Kommunikationstechnik,
Otto-von-Guericke-Universität Magdeburg

⁴Fachbereich Ingenieurwesen, Technische Fachhochschule Wildau

`heike.ruppertshofen@fh-kiel.de`

Abstract. An approach for automatic object localization in medical images utilizing an extended version of the generalized Hough transform (GHT) is presented. In our approach the shape model in the GHT is equipped with specific model point weights, which are used in the voting process. The weights are adjusted in a discriminative training procedure, which aims at a minimal localization error of the GHT. Furthermore, these weights yield information about the importance of points, such that non-discriminative points can be excluded from the model. Thereby, the size of the model is decreased, thus reducing processing time. The algorithm is presented in 2D but should be easily extendable for 3D images.

1 Introduction

The task of object localization in medical images is of high interest in many applications in the area of computer-aided diagnosis, e.g., for a fully-automatic segmentation of organs or bones.

For model-based segmentation an initial positioning of the model in the image is required. This initialization is often carried out manually or specialized solutions are developed for the particular problem at hand using, e.g., gray-value thresholding, morphological operators or anatomical knowledge [1]. More advanced, automatic approaches are given by an atlas-based registration [2] or an initial global search of the object conducted, e.g., with evolutionary algorithms [3], template matching [4] or the generalized Hough transform (GHT) [5].

In this work, we will focus on the GHT, which, in general, is a robust method and therefore well suited even for the localization of partially occluded or noisy objects. However, one drawback is the high computational demand given by the vast global search, as it is the case for all aforementioned automatic procedures. To reduce its computational complexity two measures have been undertaken. On the one hand, we will not search for scaled or rotated occurrences of our object;

but instead, will try to include this variation into our model. On the other hand, a weight will be assigned to each model point relative to its importance such that nonrelevant points can be removed from the model; thus reducing model size and runtime. Model point weights are determined through a log-linear combination of model point knowledge and a discriminative training [6] with respect to a minimal localization error.

2 Materials and Methods

2.1 Generalized Hough Transform

The GHT [7] is a model-based method for object localization utilizing a voting process to identify the likeliest position of the object of interest. To this end, the image space is mapped to a transformation parameter space, known as Hough space, which depicts possible locations of the object.

The shape model of the object is represented by vectors from each model point to a given reference point, usually the center of the model. Furthermore, directional information obtained from the gradient direction is added to each point. This information is stored in a look-up-table, the so called r-table, where model points with similar gradient directions are grouped together.

To perform the localization, an edge image is created from the original image. For each edge point e_i the model points m_j with similar gradient direction are retrieved from the r-table. Through the relation $c = e_i - m_j$ the position c of a possible reference point is determined and the vote in the corresponding cell in the Hough space is increased by the weight of the model point m_j .

At the end of the voting process the Hough space is searched and the cell with the highest vote is returned as the assumed position of the object.

The GHT can be extended to find rotated or scaled occurrences of the object of interest as well. However, due to a vast rise in complexity these extensions will not be considered here.

The model used in our extended GHT is built by extracting a number of edge points from a predefined region of interest around the reference point, which avoids the manual generation of a shape model. Yet, the model may contain points, which are not relevant for object localization. To reduce its size, the importance of model points is determined using a discriminative training algorithm as described below and points with low importance are eliminated thereafter.

2.2 Discriminative Training of Model Point Weights

Let us regard the model points of our initial model as individual sources of knowledge via their contribution to the Hough space. The aim of the training approach is to determine the importance of points which is achieved by an adequate combination of knowledge sources and a minimal localization error criterion.

For the derivation of the training approach, we have to take a probabilistic view of the GHT. Instead of searching for the maximal number of votes in the

Hough space, we define a posterior probability

$$p(c_i|X) = \frac{N_i}{N} \quad (1)$$

with c_i being a cell i in the Hough space, X a set of features extracted from a given image and N_i , N being the number of votes in cell i and in the complete Hough space. By identifying the cell with the highest posterior probability the localization task is now realized with a Bayesian classifier $\hat{c} = \arg \max_{c_i} p(c_i|X)$.

The posterior probability of the Hough cells can be further divided into contributions from the individual model points

$$p_j(c_i|X) = \frac{N_{i,j}}{N_j} \quad (2)$$

where $N_{i,j}$ and N_j are the number of votes cast by model point j . A recombination of the model point posteriors can be achieved through a log-linear modeling following the maximum entropy principle [8]

$$p_\Lambda(c_i|X) = \frac{\exp\left(\sum_j \lambda_j \cdot \log p_j(c_i|X)\right)}{\sum_k \exp\left(\sum_j \lambda_j \cdot \log p_j(c_k|X)\right)} \quad (3)$$

The coefficients $\Lambda = \{\lambda_j\}_j$ are the model point weights, which need to be estimated. To this end an error function E is defined, which accumulates the localization error on the different images X_n

$$E(\Lambda) = \sum_n \sum_i \varepsilon(\tilde{c}_n, c_i) \cdot \frac{p_\Lambda(c_i|X_n)^\eta}{\sum_k p_\Lambda(c_k|X_n)^\eta} \quad (4)$$

Here, the Euclidian distance of the correct cell \tilde{c}_n to a cell in the Hough space c_i is chosen as error measure ε weighted by an indicator function. The exponent η in the indicator function controls the influence of the alternative hypotheses c_k on the error measure. By applying a gradient descent scheme to (4) the optimal weights λ_j can be determined with respect to a minimal localization error [6].

2.3 Material and Design of Experiments

The algorithm was tested on a set of 30 thorax radiographs, which were acquired in a follow-up study of 11 patients. The images have an isotropic resolution of 0.185 mm with varying image sizes. To evade noise in the training process, the images were downsampled twice with a Gaussian filter to a spacing of 0.74 mm.

The given task is to localize the collar bone in all images. For this purpose the intersection of collar bone and lung wall was annotated to obtain a ground-truth for training as well as evaluation of the algorithm.

The dataset was divided into a training and a test dataset, each consisting of 15 images. The model used for the localization task was created from a set of 8 training images.

Table 1. The table shows a comparison of the initial and the final model regarding the number of model points, the average runtime per image in minutes and the localization error on the training and test data in mm displayed as "mean \pm std (max)".

	model size	runtime	error on training data	error on test data
initial model	2684	1.2	2.2 \pm 2.1 (6.9)	3.8 \pm 2.6 (10.2)
final model	920	0.7	1.9 \pm 2.1 (8.2)	3.0 \pm 2.1 (6.7)

Using this model the GHT is performed on the training images with equal weights for all points, followed by the computation of the optimal model point weights as described above. Based on these weights a subset of points necessary for object localization is determined as final model. Subsequently, the weighted final model is tested on the complete dataset.

Experiments were run on a desktop pc with 2.4 GHz. The GHT is implemented in Matlab and not yet optimized for speed.

3 Results

The initial model extracted from part of the training images and the final model resulting after the discriminative training are shown in Fig. 1. It is clearly visible that the computed point weights allow a strong reduction of the model size. Only about 35% of the initial points were kept, reducing the processing time by almost 50% as can be seen in Tab. 1.

By further examining the model it becomes obvious that not only the edge of the collar bone is important for its detection; but also the lung wall and surrounding bones are of high interest. The inclusion of these structures helps to distinguish the collar bone from the ribs, which show a similar edge response.

Since the initial model was extracted from the training data a very good localization result was achieved on these images, which even grew slightly better for the final model. As one could expect the error on the unknown test data is slightly higher, but should be sufficient for the initialization of a segmentation.

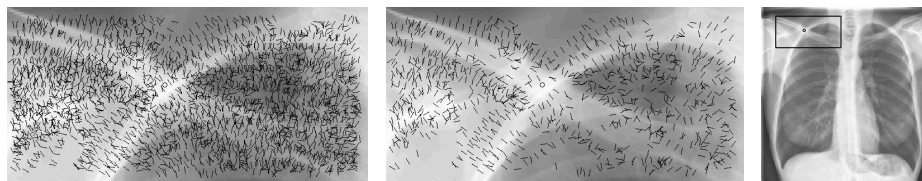


Fig. 1. Comparison of the initial model (left) extracted from the images and the final model (middle). Shown are the point locations with corresponding gradient directions. For illustration an annotated example image is displayed right

4 Discussion

The combination of the GHT with the discriminative training algorithm to determine adequate model point weights shows promising results. The introduction of point weights into the model brings about several advantages. First, no initial shape model needs to be extracted from the data via automatic or manual segmentation, but a number of edge points from the region of interest are sufficient for the GHT algorithm. Second, large models can be exploited during training, since the number of model points can be diminished afterwards keeping only discriminative points. In addition, neighboring structures, which may be helpful in the localization task, can easily be included in the model by extracting a larger ROI around the object of interest. Furthermore, by extracting the model directly from the dataset, the variation of the object of interest in the underlying images is captured. Thus no rotation or scaling of the model was necessary in the process of the GHT. This implies a considerable reduction of runtime and memory need. Runtime was further decreased through the reduction of model size, while the localization error slightly improved to about 3.0 mm on average.

The algorithm is expected to be applicable to detect arbitrary objects, which show a clear response in the edge image. Experiments will be conducted to reveal further possible features to be able to detect objects, which are not characterized by their edge image. Future work will also include an iterative approach, where random points are gradually included in the model and evaluated by DMC to further improve localization accuracy.

Acknowledgement. We thank Dr. Cornelia Schaefer-Prokop, AMC Amsterdam for the abundance of radiographs.

References

1. Heimann T, van Ginneken B, Styner M, et al. Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Trans Med Imaging*. 2009;28(8):1251–65.
2. Seghers D, Slagmolen P, Lambelin Y, et al. Landmark based liver segmentation using local shape and local intensity models. In: *Proc MICCAI*; 2007. p. 135–42.
3. Heimann T, Münzinger S, Meinzer HP, et al. A shape-guided deformable model with evolutionary algorithm initialization for 3D soft tissue segmentation. In: *Proc IPMI*; 2007. p. 1–12.
4. Barthel D, von Berg J. Robust automatic lung field segmentation on digital chest radiographs. *Int J CARS*. 2009;4(Suppl 1):326–7.
5. Schramm H, Ecabert O, Peters J, et al. Towards fully automatic object detection and segmentation. In: *SPIE*; 2006. p. 614402.
6. Beyerlein P. Discriminative model combination. In: *Proc ICASSP*; 1998. p. 481–4.
7. Ballard DH. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognit*. 1981;13(2):111–22.
8. Jaynes ET. Information theory and statistical mechanics. *Phys Rev*. 1957;106(4):620–30.

Statistische 3D Formmodellierung mittels quasi-verzerrungsfreier sphärischer Parametrisierung

Sebastian T. Gollmer, Thorsten M. Buzug

Institut für Medizintechnik, Universität zu Lübeck
gollmer@imt.uni-luebeck.de

Kurzfassung. In diesem Beitrag wird die Integration einer etablierten Methode für die Parametrisierung geschlossener Oberflächen mit minimaler Verzerrung in einen effizienten Algorithmus für die Erstellung statistischer Formmodelle (SFM) vorgestellt. Dieser bestimmt die erforderlichen Punkt-Korrespondenzen (Landmarken) durch Minimierung der sog. Description Length des SFM. Da diese Optimierung in einer Parameterdomäne erfolgt, ist zunächst eine entsprechende Initialisierung in Form von parametrisierten Formrepräsentationen erforderlich. Unvermeidbare geometrische Verzerrungen in der Parameterdomäne müssen möglichst gering sein, da das resultierende SFM die eigentliche Objektform sonst nur unzureichend repräsentiert. Während vorherige Arbeiten eine Maximierung der Flächentreue der initialen Parametrisierungen oder eine geeignete Umverteilung der Landmarken vornehmen, erlaubt das von uns vorgestellte Verfahren die Minimierung der Flächenverzerrung unter gleichzeitig bestmöglicher Erhaltung der Winkelverhältnisse.

1 Einleitung

Statistische Formmodelle (SFM) [1] integrieren die zu erwartende Formvariabilität eines bestimmten Objekts auf der Basis einer repräsentativen Trainingspopulation. Die elastische Adaption des Modells an patientenspezifische Formausprägungen erfolgt somit a priori innerhalb plausibler Grenzen. Bei der Erstellung des SFM stellt die Identifikation von Punkt-Korrespondenzen (Landmarken) über alle Instanzen der Trainingspopulation die wesentliche Herausforderung dar. Für 3D-Modelle sind ausschließlich automatisierte Verfahren praktikabel. Eine Möglichkeit hierzu bietet die Definition inhärenter Korrespondenzen durch geeignete Parametrisierung der Forminstanzen, wobei sich für zur 2-Sphäre Ω^3 homeomorphe Objekte, sphärische Harmonische (SPHARM-Modell [2]) besonderes eignen.

Am erfolgreichsten lässt sich die Korrespondenzfindung als Optimierungsproblem formulieren, mit dem Ziel, diejenigen Korrespondenzen zu finden, welche das „beste“ SFM im Sinne eines bestimmten Gütekriteriums repräsentieren. Als Goldstandard für die Zielfunktion hat sich die Minimierung der sog. *Description Length* (DL) des Modells [3] etabliert und wurde im Hinblick auf Effizienz weiterentwickelt [4]. In der Praxis lässt sich die Optimierung nur in einer geeigneten Parameterdomäne lösen, weshalb zunächst die initiale Parametrisierung der

Trainingsinstanzen erforderlich ist. Dabei auftretende unvermeidbare Verzerrungen in der Parameterdomäne müssen möglichst gering sein, da das resultierende SFM die eigentliche Objektform sonst nur unzureichend repräsentiert. Aus diesem Grund minimieren Davies et al. [5] die Flächenverzerrung über alle Parametrisierungen der Trainingspopulation. Um dagegen die ausgeprägten Flächenverzerrungen der in [4] verwendeten konformen Parametrisierung zu kompensieren, schlagen Heimann et al. [6] eine geeignete Umverteilung der Landmarken (Remeshing) während und bzw. oder nach der Optimierung vor.

Der wesentliche Beitrag der vorliegenden Arbeit ist die Implementierung einer effizienten Korrespondenzfindung [4, 6] unter Verwendung des kompletten Parametrisierungsverfahren nach [2], während in [3, 5] lediglich die Initialisierung aus [2] verwendet wird. Wir erhalten dadurch Parametrisierungen mit minimaler Flächenverzerrung und bestmöglicher Winkeltreue. Ein Remeshing wie in [6] wird damit obsolet und die dabei induzierte, potenziell negative Beeinflussung der Optimierung vermieden.

2 Material und Methoden

2.1 Statistische Formmodelle

Die Formen $S_i \subset \mathbb{R}^d$, $i = 1, \dots, n_s$ der Trainingspopulation lassen sich jeweils durch n_p korrespondierende Landmarken repräsentieren und somit mit dem Vektor $\mathbf{x}_i \in \mathbb{R}^{dn_p}$ beschreiben. Die Formvariabilität ergibt sich aus der räumlichen Variation der Landmarken über alle Formen $\{S_i\}$. Dazu werden die Durchschnittsform $\bar{\mathbf{x}}$ sowie, mittels Singulärwertzerlegung, die n_m Eigenvektoren $\{\mathbf{p}_m\}$ und -werte $\{\lambda_m\}$ der Landmarkenvariation berechnet. $\mathbf{x}^* = \bar{\mathbf{x}} + \sum_{m=1}^{n_m} \mathbf{p}_m b_m^*$ ist damit eine beliebige valide Forminstanz mit den Formparametern $\{b_m^* \in \mathbb{R}\}$.

2.2 Initialisierung und Optimierung der Punkt-Korrespondenzen

Initiale Korrespondenzen lassen sich durch äquidistante Verteilung der n_p Landmarken in einer geeigneten Parameterdomäne Ω (hier: Ω^3) und die Repräsentation der Forminstanzen $\{S_i\}$ durch die entsprechenden Parametrisierungsfunktionen $\{\omega_i\}$ bestimmen (Abb. 1). In 3D ist die Form S_i eine Oberfläche, die bijektive Abbildung $f : \Omega^3 \rightarrow S_i$ bezeichnet eine Oberflächenparametrisierung und Ω_i die Oberfläche in Ω^3 . Mit den Parameterkoordinaten $\theta \in [0, \pi]$ und $\phi \in [0, 2\pi]$ folgt $\Omega^3 \ni (\theta, \phi) \rightarrow \omega_i(\theta, \phi) = \rho \in S_i$.

Formal führt die inverse Abbildung $f^{-1} : S_i \rightarrow \Omega^3$ zur gesuchten initialen Oberfläche Ω_i , wobei sich die Polygonflächen und -winkel nicht gleichzeitig vollständig erhalten lassen. In [4, 6] wird die winkeltreue (konforme) Abbildung f_K von Gu et al. [7] verwendet, welche hier durch die verzerrungsminimierende Abbildung f_A nach Brechbühler et al. [2] ersetzt wird. Initiale Werte für θ und ϕ ergeben sich durch Lösen der Laplace'schen Gleichungen $\nabla\theta = 0$ und $\nabla\phi = 0$. Die jeweiligen Dirichlet-Randbedingungen sind zum einen $\theta_{\text{Nordpol}} = 0$ und $\theta_{\text{Südpol}} = \pi$ (Zuweisung zum Vertex mit min./max. z -Koordinate) sowie die

Periodizität von 2π und die Fixierung eines einzelnen Vertex. Anschließend wird die Flächenverzerrung mit Hilfe des Newton-Verfahrens unter Berücksichtigung von drei Nebenbedingungen minimiert: Erhaltung von Vektorlängen, Polygonflächen und Polygonwinkel. Die Oberfläche Ω_1 der ersten Forminstanz in Abb. 1(b) zeigen Abb. 1(c) für f_K und Abb. 1(d) für f_A . Gleiche Farben stellen Korrespondenzen zwischen S_1 und Ω_1 dar.

Die Re-Positionierung der Landmarken erfolgt im Sinne der Minimierung von

$$F = \sum_m \mathcal{L}_m, \text{ mit } \mathcal{L}_m = \begin{cases} 1 + \log(\lambda_m/\lambda_{\text{cut}}) & \text{für } \lambda_m \geq \lambda_{\text{cut}} \\ \lambda_m/\lambda_{\text{cut}} & \text{für } \lambda_m < \lambda_{\text{cut}} \end{cases} \quad (1)$$

wobei λ_{cut} in Abhängigkeit von der Rauschbehaftung der Trainingspopulation zu wählen ist [4] (hier: $\lambda_{\text{cut}} = (0.3/\bar{r})^2$, mittlerer Radius der $\{S_i\}$: $\bar{r} = 50$ Voxel). (1) ist eine Approximation der DL des SFM mit dem Vorteil einer semi-analytisch bestimmbaren Ableitung. Die Ableitung der Re-Parametrisierung $\Phi_i(\theta, \phi)$ und somit die Gradientenrichtung auf Ω_i wird mittels finiter Differenz unter Verwendung eines lokalen Re-Parametrisierungsverfahrens [4] bestimmt.

2.3 Evaluierung der Modellgüte

Wir quantifizieren den Einfluss der initialen Parametrisierung auf die Modellgüte durch Vergleich von jeweils zwei, unter Verwendung von f_K und f_A erstellten SFM der Hippocampus-Region des menschlichen Gehirns. Dabei kommen die bekannten Gütekriterien Spezifität \mathcal{S} und Generalisierungsfähigkeit \mathcal{G} zum Einsatz. Bei einem spezifischen SFM weisen N (hier: $N = 1000$) unter zufälliger Variation von $\{b_m^*\}$ erstellte Rekonstruktionen S^* eine große Ähnlichkeit mit den Trainingsinstanzen $\{S_i\}$ auf. Die Fähigkeit zur Abbildung unbekannter Forminstanzen wird mittels Leave-one-out Tests bewertet und als \mathcal{G} bezeichnet.

Generell hängen \mathcal{S} und \mathcal{G} von der Anzahl n_m der berücksichtigten Parameter $\{b_m^*\}$ ab. Für die Bestimmung der Ähnlichkeit von S^* zu den $\{S_i\}$ sind verschiedene Distanzmaße verwendbar. Ein anerkannter und effizienter Ansatz ist die Berechnung des mittleren absoluten Abstands der Korrespondenzen. Die Formähnlichkeit des Modells mit der Trainingspopulation und damit der Objektklasse

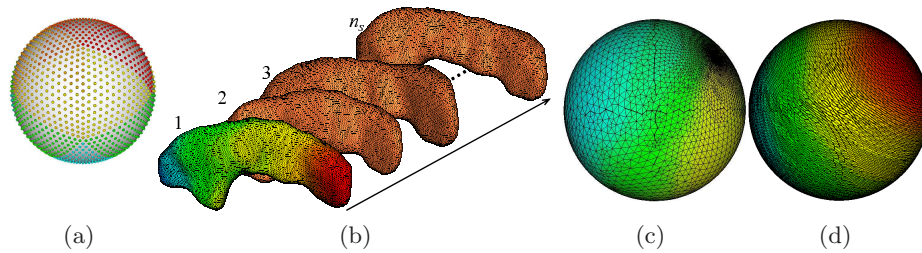


Abb. 1. Initialisierung der Punktkorrespondenzen: Verteilung der äquidistanten Landmarken (a) über die Forminstanzen (b) unter Verwendung der Parametrisierungen $\{\omega_i\}$. (c) und (d) zeigen Ω_1 unter Verwendung von f_K bzw. f_A .

bleibt dabei jedoch unberücksichtigt. Somit ist die Quantifizierung unzureichender Formrepräsentationen, z.B. infolge parametrisierungsbedingter Verzerrungen, auf diese Weise nicht möglich. Aus diesem Grund verwenden Heimann et al. [6] die Jaccard-Metrik, während wir die Oberflächen der Forminstanzen unter Verwendung des symmetrischen quadratischen Abstandes vergleichen.

3 Ergebnisse

Qualitative und quantitative Ergebnisse sind in Abb. 2 dargestellt. In Abb. 2(a) sind zwei Hippocampus-Modelle ($n_s = 21, n_p = 2562$) unter cranialer Ansicht zu sehen, die unter Verwendung von f_K (oben) und f_A (unten) erstellt wurden. Die Landmarken-Verteilungen \bar{x} sind den dadurch repräsentierten Formen \bar{S} jeweils überlagert. Die quantitative Auswertung von \mathcal{S} und \mathcal{G} dieser, sowie von zwei weiteren Modellen mit $n_p = 10242$, zeigen Abb. 2(b, c). Abgebildet sind \mathcal{S} und \mathcal{G} für Rekonstruktionen S^* unter Berücksichtigung von bis zu zehn Hauptmoden, was für alle Modelle einer Gesamtvarianz von $\geq 90\%$ entspricht.

4 Diskussion

Aus den Ergebnissen in Abb. 2 ist die starke Abhängigkeit der Modellgüte von der initialen Parametrisierung unmittelbar ersichtlich. Bereits bei Betrachtung der exemplarisch für die Form S_1 dargestellten, mittels f_K bzw. f_A erstellten Parametrisierungen ω_1 in Abb. 1(c, d) liegt die Vermutung nahe, dass die unter f_K insbesondere an den Regionen von S_1 mit der größten Krümmung auftretenden, ausgeprägten Flächenverzerrungen zu einer entsprechend defizitären Repräsentation dieser Regionen durch die Landmarken führt. Dementsprechend stellt die klar erkennbare schlechte Güte des Modells in Abb. 2(a) (oben), sowie Fehler von mehreren mm bei der Auswertung von \mathcal{S} und \mathcal{G} (Abb. 2(b, c)) eine logische Konsequenz dar. Die exzessive Abtastung von Ω^3 mit $n_p = 10242$ Landmarken scheint intuitiv eine Möglichkeit darzustellen, der Unterrepräsentation bestimmter Regionen des Objektes entgegenzuwirken. Gemäß Abb. 2(b, c) ist auf diese Weise eine Reduzierung der Fehler für \mathcal{S} und \mathcal{G} um einen Faktor von ca. 4

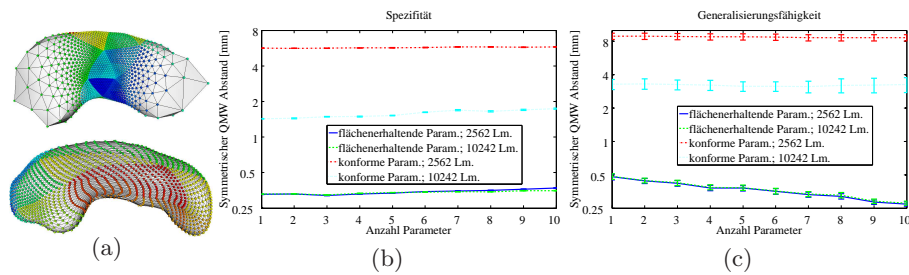
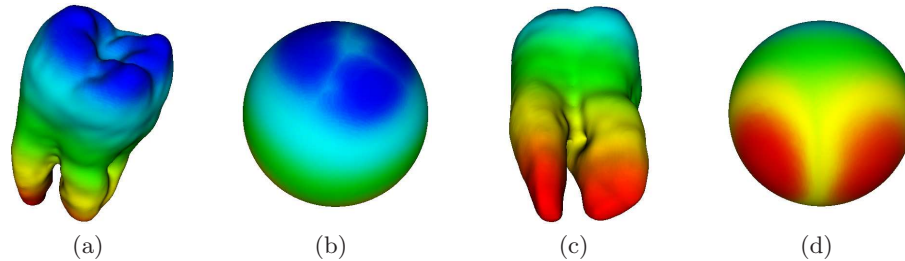


Abb. 2. Landmarken-Repräsentationen \bar{x} (a) bei Initialisierung mit f_K (oben) und f_A (unten) sowie zugehörige Gütekriterien \mathcal{S} (b) und \mathcal{G} (c).

Abb. 3. Parametrisierung der komplexen Anatomie eines Backenzahns (The Volume Library, <http://www9.informatik.uni-erlangen.de/External/vollib/>) mittels f_A .



bzw. 2,5 gegenüber dem SFM mit $n_p = 2562$ möglich. Der Unterschied zu den basierend auf f_A erstellten SFM liegt jedoch weiterhin im Bereich einer Größenordnung, so dass die Vergrößerung der Anzahl der Landmarken keine adäquate Kompensationsmethode darstellt.

Die Ergebnisse zeigen, dass wir unter Verwendung von f_A in der Lage sind, optimale, die Objektklasse valide repräsentierende SFM unter Verwendung eines effizienten Algorithmus zu erstellen. Die Entbehrlichkeit eines *Remeshing* während der Optimierung stellt nicht nur eben diese Effizienz sicher, sondern vermeidet auch eine Verringerung der Stabilität der Konvergenz zum globalen Minimum. Wir wollen die vorgestellte, erfolgreiche Kombination ebenfalls für die Erstellung von SFM für Anatomien mit komplexer Geometrie verwenden. Die Plausibilität der Parametrisierung des Backenzahns in Abb. 3(a, c) weist darauf hin, dass f_A hierzu das nötige Potenzial liefert (Abb. 3(b, d)).

Danksagung. Wir danken M. Styner und Mitarbeitern, UNC Neuro Image Laboratories, für die zur Verfügung Stellung der Hippocampus-Daten.

Literaturverzeichnis

1. Cootes TF, Taylor CJ, Cooper DH, et al. Active shape models: their training and application. *Comput Vis Image Underst.* 1995;61(1):38–59.
2. Brechbühler C, Gerig G, Kübler O. Parametrization of closed surfaces for 3-D shape description. *Comput Vis Image Underst.* 1995;61(2):154–70.
3. Davies RH, Twining CJ, Cootes TF, et al. 3D statistical shape models using direct optimisation of description length. In: *Proc ECCV*; 2002. p. 1–17.
4. Heimann T, Wolf I, Williams T, et al. 3D active shape models using gradient descent optimization of description length. In: *Proc IPMI*; 2005. p. 566–77.
5. Davies RH, Twining CJ, Taylor CJ. Consistent spherical parameterisation for statistical shape modelling. In: *Proc ISBI*; 2006. p. 1388–91.
6. Heimann T, Wolf I, Meinzer HP. Automatic generation of 3D statistical shape models with optimal landmark distributions. *Methods Inf Med.* 2007;46(3):275–81.
7. Gu X, Wang Y, Chan TF, et al. Genus zero surface conformal mapping and its application to brain surface mapping. *IEEE Trans Med Imaging.* 2004;23(8):949–58.

3D Statistical Shape Model Building using Consistent Parameterization

Matthias Kirschner, Stefan Wesarg

Graphisch Interaktive Systeme, TU Darmstadt
matthias.kirschner@gris.tu-darmstadt.de

Abstract. We propose a new correspondence optimization algorithm for building 3D statistical shape models (SSMs) of genus-0 shapes. The main contribution of our work is the use of parameter space propagation to generate consistent spherical parameterizations of the training shapes. We present evaluation results for two data sets: A set of 30 liver shapes from different patients, and a set of 25 left ventricles covering the cardiac cycle of a single patient. Our evaluation shows that the use of parameter space propagation improves the robustness of correspondence optimization algorithms and leads to fast convergence.

1 Introduction

In medical imaging, statistical shape models (SSMs) are mainly employed in segmentation algorithms like the active shape model (ASM), which shows good performance in terms of both segmentation accuracy and running time [1]. SSMs are learned from a set of training examples, which have to be provided in a landmark representation, that is, each training shape has to be represented with the same number of points, and points with same index on different shapes have to describe the same anatomical feature. The challenging optimization problem of computing a landmark representation automatically from a set of unnormalized training meshes is known as the correspondence problem.

To tackle the correspondence problem, Kotcheff and Taylor [2] introduced a general optimization scheme, which we follow in our work: Each training shape is mapped to a suitable parameter space, and correspondence is then optimized by reparameterization. The correspondence of the reparameterized shapes can be assessed with various objective functions, for example the Minimum Description Length (MDL) objective function [3].

The natural parameter space for the construction of 3D SSMs of shapes with genus-0 topology is the unit sphere. Davies et al. [4] construct area-preserving spherical parameterizations of the training shapes, whereas Heimann et al. [5] use conformal mapping to generate angle-preserving spherical parameterizations of liver shapes. We concentrate on area-preserving parameterizations, because they allow for an easy reconstruction of the shapes using uniform sampling. It is crucial to parameterize the shapes consistently – that means to map the same anatomical features to similar regions in the parameter spaces – in order

to decrease the convergence time and to prevent that the optimization process gets trapped in a poor local optimum [6]. Recently, we introduced parameter space propagation [7] to generate consistent parameterizations more efficiently and without a complex optimization process as proposed in [6, 8]. The key idea of our method is to propagate the parameterization of a reference shape to all other shapes.

In this paper, we propose a new optimization algorithm that automatically establishes correspondence of shapes with genus-0 topology. The main novelty is the use of parameter space propagation in order to generate consistent parameterizations. Our evaluation shows that the use of parameter space propagation avoids that correspondence optimization algorithms get trapped in poor optima and leads to a fast convergence towards a good solution.

2 Materials and Methods

2.1 Consistent Parameterization by Propagation

We use parameter space propagation [7] in order to consistently generate spherical parameterizations of the training shapes. In this method, an area-preserving parameterization of an arbitrarily selected reference shape S_{ref} is computed. Every other shape S is aligned to the reference shape using the ICP algorithm in order to derive a common coordinate system. Based on the Euclidean distance in this coordinate system, a fuzzy correspondence between points in S and S_{ref} is established. The points of S are then mapped to the unit sphere by interpolation of the parameter space coordinates of the corresponding points of S_{ref} . A subsequent correction method handles triangles that are inverted or overlap on the unit sphere. Parameter space propagation is a heuristic method, but works robustly for typical organ shapes. By computing surface normals of the triangles of the generated parameterizations, we verified that only valid parameterizations were generated for our test data sets.

2.2 Optimization Algorithm

In this section, we present our optimization algorithm, which minimizes the objective function $\mathcal{L} = \sum_{i=1}^t \log(\lambda_i + \epsilon)$, where the λ_i are the eigenvalues of the $t \times t$ covariance matrix of the shapes, which is determined by numerical integration [3, 8]. We set the regularization constant ϵ to 0.005. The correspondence measure \mathcal{L} is (up to a constant term) the so-called DetCov function [2].

Our algorithm follows the general minimization approach with parametric regularization [8]. We start with generating sampling points for each shape which are initially uniformly distributed over the sphere. The optimization works iteratively: In each iteration, we select a shape uniformly at random and re-parameterize its sampling points using two kinds of re-parameterizations:

- *Uniform rotation:* We consider rotations of the sampling points around the x, y and z-axis, one after another, in randomly determined order. For each

axis, we choose the optimal rotation angle between -15° and $+15^\circ$ using a line search algorithm.

- *Clamped plate spline warps*: We map the sampling points of a randomly selected spherical cap to the unit circle using orthographic projection, manipulate them using a clamped plate spline warp [8] and project them back to the sphere. The optimal parameters for the spline transformation are found using gradient descent optimization, with numerically estimated gradients.

Additionally, we optimize the pose parameter rotation directly. The other pose parameters are handled prior to the optimization process, by scaling the shapes to same size and translating them to the origin.

2.3 Evaluation Method

With our evaluation, we want to show that 1) our optimization algorithm is able to generate SSMs for different sets of genus-0 shapes and that 2) the use of parameter space propagation enables us to compute SSMs of high-quality. To show 1), we tested our algorithm on data sets with different organs: A set of 30 liver shapes from different patients, and a set of 25 left ventricles (LVs) which cover the whole cardiac cycle of a single patient with a myocardial infarction. To verify our second hypothesis, we started our algorithm with three different kinds of parameterizations: Parameterizations generated with our propagation method, parameterizations generated using Davies’s method (see [7] for implementation details) and parameterizations which are generated independently from each other. We used a fixed number of iterations ($n = 3000$). Prior to the parameterization, the liver shapes are aligned into a common coordinate system, to ensure

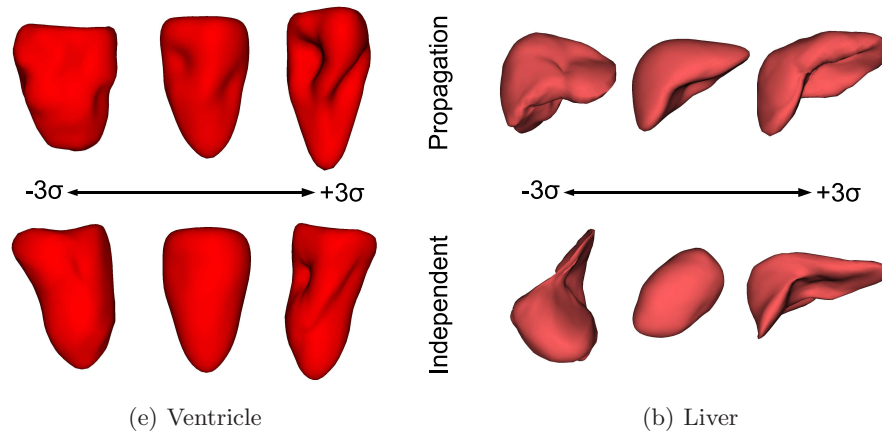


Fig. 1. The first mode of variation of the ventricle and liver models, optimized using parameterizations generated with our method (top row) and independently computed parameterizations (bottom row). $\sigma = \sqrt{\lambda_1}$ denotes the standard deviation of the first mode of the respective SSM.

that our results are not tampered by poor initial alignment. The meshes in the LV data set are already aligned, as they were generated from a single 4D MRI volume. In order to compare the quality of the generated models, we calculate the standard measures specificity and generalization ability as described in [8], by sampling 10000 shapes randomly from the respective probability distribution.

3 Results

Figure 2 shows the evolution of the objective function during the execution of the algorithm. The figure shows that the algorithm converges quickly towards a local optimum in case of consistent parameterization. But the final objective function value after 3000 iterations is smaller in case of independently computed parameterizations. However, only the shapes of the models generated with consistent parameterization look plausible, as illustrated in Fig. 1. This visual impression is confirmed by evaluating the models using the specificity and generalization ability measures. On both data sets, our optimization algorithms produces significantly better models in case of consistent parameterization. A comparison between the two different methods for consistent parameterization shows that the parameter space propagation generates better models on the ventricle data set (Fig. 3). On the liver data set, the differences between the two approaches are very small. Here, initialization with parameter space propagation leads to models with slightly better generalization ability, but which are also slightly less specific than the models generated using initialization with Davies’s method. The required optimization time for a liver model was approximately 300 minutes, and 210 minutes for a ventricle model on a 2.4 GHz CPU.

4 Discussion

We presented a new algorithm for correspondence optimization, which employs parameter space propagation to generate consistent spherical parameterizations.

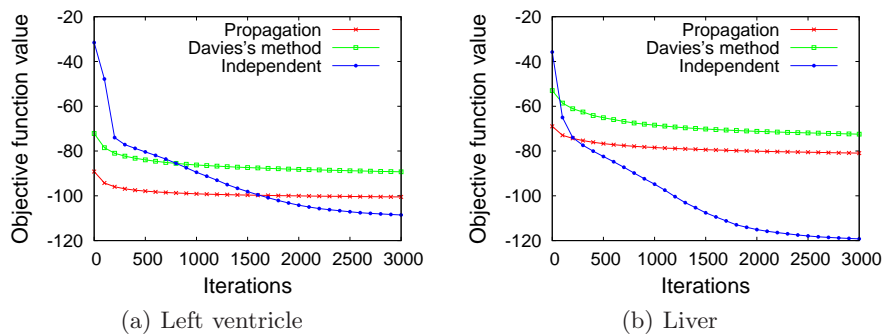
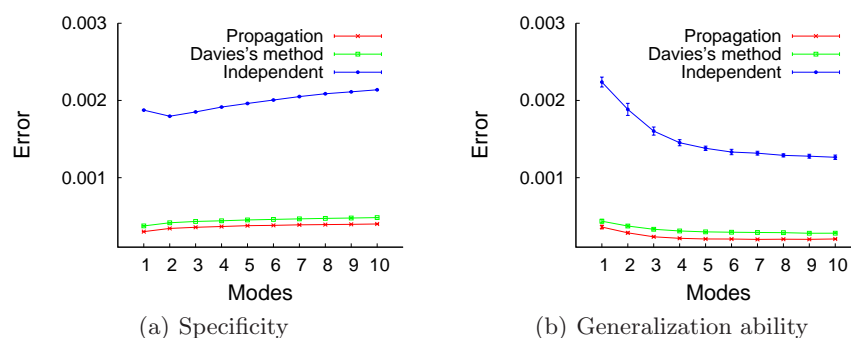


Fig. 2. The evolution of the objective function during execution of the algorithm using different methods of parameterization.

Fig. 3. Specificity and generalization ability of the generated ventricle models.

Using this technique, the algorithm quickly converges to a SSM which captures the variation of organ shapes present in the training set well. Our evaluation shows that the use of parameter space propagation improves the robustness of the correspondence optimization process. In comparison to Davies's method [6], initialization with parameter space propagation produces models which are more compact, but have comparable if not better quality in terms of specificity and generalization ability.

Possible directions for future research are an extended evaluation of our approach on other organ shapes, a comparison of different objective functions as well as algorithmic techniques to decrease the computation time of the algorithm.

References

1. Heimann T, van Ginneken B, Styner M, et al. Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Trans Med Imaging*. 2009;28:1251–65.
2. Kotcheff ACW, Taylor CJ. Automatic construction of eigenshape models by direct optimization. *Med Image Anal*. 1998;2:303–14.
3. Davies RH, Twining CJ, Cootes TF, et al. A minimum description length approach to statistical shape modeling. *IEEE Trans Med Imaging*. 2002;21(5):525–37.
4. Davies RH, Twining CJ, Cootes TF, et al. 3D statistical shape models using direct optimisation of description length. In: *Proc ECCV*; 2002. p. 3–20.
5. Heimann T, Wolf I, Williams TG, et al. 3D Active shape models using gradient descent optimization of description length. In: *Proc IPMI*; 2005. p. 566–77.
6. Davies RH, Twining CJ, Taylor CJ. Consistent spherical parameterisation for statistical shape modelling. In: *Proc ISBI*; 2006. p. 1388–91.
7. Kirschner M, Wesarg S. Construction of groupwise consistent shape parameterizations by propagation. In: *Proc SPIE*; 2010. p. to appear.
8. Davies RH, Twining CJ, Taylor CJ. *Statistical Models of Shape: Optimization and Evaluation*. Springer; 2008.

Pharmakokinetische Modellierung von FMISO-PET/CT Bildgebung bei Plattenepithelkarzinomen im Kopf-Halsbereich

Jens-Christoph Georgi¹, Rolf Bippus¹, Wenli W. Wang³, Nancy Y. Lee⁴,
Manoj Narayanan², Heiko Schöder⁵, Jose Guillem⁶, John L. Humm³

¹Philips Research Europe, Molecular Imaging Systems, Aachen

²Philips Research North America, Briarcliff Manor, NY, USA

Abteilung/Klinik für ³Medizinische Physik, ⁴Strahlentherapie, ⁵Radiologie, ⁶Chirurgie
Memorial Sloan-Kettering Cancer Center, New York, NY, USA

jens.christoph.georgi@philips.com

Kurzfassung. Ziel der Arbeit war es die Möglichkeiten der bildgestützten Hypoxiequantifizierung für Kopf-Hals-Tumore auszuloten. Dazu wurde die Korrelation pharmakokinetischer Parameter aus dynamischer FMISO PET Bildgebung mit Messungen auf statischen Bildern untersucht. Folgende Parameter eines 2-Gewebekompartiment-Modells wurden mit Gewebe/Blut-Verhältnisse (TBR) im statischen Bild korreliert: metabolische Rate K_i , fraktionelles Blutvolumen β und lokale Perfusion k_1 . Ausgewertet wurden 9 Patienten mit Diagnose auf lokalisiertes Plattenepithelkarzinomen im Kopf-Halsbereich ohne vorherige Chemo- oder Radiotherapie. Die Registrierung der Bilddaten und die anschließenden pharmakokinetischen Analysen wurde mit Hilfe der VoxulusTM Software der Philips Forschung durchgeführt, wie sie auf der Philips ImalyticsTM Workstation zur Verfügung steht. Für Tumor ROIs konnte eine starke positive Korrelation zwischen mittlerem K_i und TBR gezeigt werden. Ebenso eine leichte negative Korrelation zwischen k_1 and K_i und eine moderat positive Korrelation zwischen β and K_i , jedoch keine Korrelation zwischen β und k_1 . Die Ergebnisse legen nahe, dass die Verwendung des statischen Gewebe/ Blut-Verhältnisses alleine unzureichend ist um Hypoxie im Tumorgewebe eindeutig zu identifizieren. Ein robusteres Kriterium verspricht die Verwendung einer Kombination der beiden Parameter TBR und K_i .

1 Einleitung

Lokale Sauerstoffunterversorgung (Hypoxie) ist aus der Literatur als negativer prognostischer Faktor für eine Vielzahl von Tumorerkrankungen bekannt: Gebärmutterhalskrebs, Sarkome sowie Prostatakarzinome. Auch bei Plattenepithelkarzinomen im Kopf-Halsbereich wurde Hypoxie als entscheidender Faktor zu Prognose des Therapieverlaufs identifiziert [1].

Das Wissen um lokale Hypoxie kann somit von entscheidender Bedeutung bei der Therapieselektion und -planung und -prognose sein. Nichtinvasive, quantitative und räumlich aufgelöste Verfahren wie die bildgestützte pharmakokinetische Modellierung versprechen für die Zukunft entsprechende Möglichkeiten.

2 Material und Methoden

2.1 Pharmakokinetische Analyse

Die pharmakokinetischen Analysen wurde mit Hilfe der VoxulusTM Software der Philips Forschung durchgeführt, wie sie auf der Philips ImalyticsTM Workstation zur Verfügung steht. Abb. 1 zeigt das verwendete Kompartimentmodell. Dieses besteht aus dem Plasmavolumen sowie zwei Gewebekompartimenten, einem reversibel und einem irreversibel bindenden Kompartiment [2].

Die entsprechenden Differentialgleichungen

$$\frac{dC_1(t)}{dt} = k_1 C_p(t) - (k_2 + k_3) C_1(t) ; \quad \frac{dC_2(t)}{dt} = k_3 C_1(t) \quad (1)$$

und die gemessene FMISO-Konzentration bestimmt sich zu

$$C_{\text{ROI}}(t) = \beta C_p(t) + C_1(t) + C_2(t) \quad (2)$$

wobei β die Vaskularisierung des Tumor Plasmakompartiments relativ zur arteriellen Eingangsfunktion ist. Das gezeigte Modell und die verwendete Software wurden zuvor ausführlich an simulierten Daten auf Reproduzierbarkeit sowie Empfindlichkeit gegenüber unterschiedlichen Störeinflüssen untersucht [3].

2.2 Patientenprotokoll und Auswertung

Ausgewertet wurden neun Patienten mit Diagnose auf lokalisiertes Plattenepithelkarzinom im Kopf-Halsbereich ohne vorherige Chemo- oder Radiotherapie. Alle Patienten wurden zwei PET/CT Studien unterworfen. Zunächst einem statischen FDG PET/CT zur Bildgebung des eigentlichen Tumors. Drei Tage später folgte dann die dynamische FMISO PET/CT Studie gemäß des in Abb. 2 gezeigten Protokolls [4]. Die statische FDG PET/CT Aufnahme wurde zur Delineation des eigentlichen Tumors (Tumor-ROI) herangezogen.

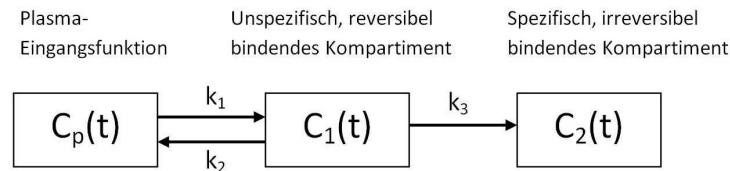
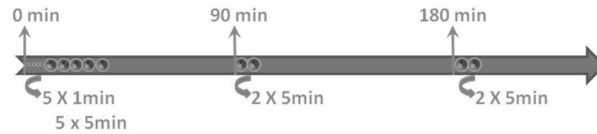


Abb. 1. Verwendetes Kompartimentmodell für FMISO in Tumorgewebe.

Abb. 2. Aquisitionsprotokoll für das dynamische FMISO PET/CT.



Auf den zuvor rigide registrierten dynamischen FMISO Daten wurden folgende Parameter individuell für jedes Voxel bestimmt, die ihren entsprechenden physiologischen Parametern zuzuordnen sind:

k_1 = lokale Perfusion

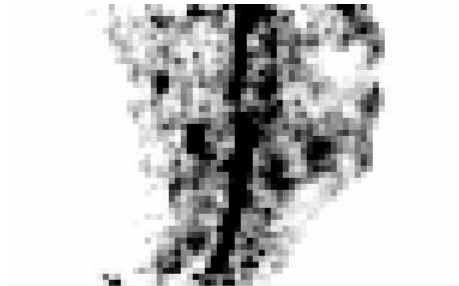
β = fraktionelles Blutvolumen

$K_i = \frac{k_1 k_3}{k_2 + k_3}$ = metabolische Rate, hypothetisiertes Maß für Hypoxie

TBR = Gewebe/Blut Konzentrationsverhältnis
zum Zeitpunkt $t = 3$ h nach Injektion

Die Schätzung der arteriellen Eingangsfunktion erfolgte auf den dynamischen FMISO-PET Bildern der Karotiden (Abb. 3).

Abb. 3. Karotide im FMISO PET zur Schätzung der arteriellen Eingangsfunktion.



3 Ergebnisse

Die Tabelle 1 zeigt die mittleren Korrelationskoeffizienten zwischen den berechneten Parametern für alle neun Patienten. Für das Verhältnis K_i zu TBR sind in Abb. 4 die mittleren Werte der einzelnen Patienten dargestellt.

Daneben wurden individuelle Scatterplots für jeden Patienten erstellt, wie sie für drei Patienten in Abbildung 5 zu sehen sind. Jeder schwarze Punkt entspricht hier einem Voxel des gesamten Bildausschnitts, die grau markierten denen der Tumor-ROI. Daneben sind jeweils im Schnittbild die dazugehörige Tumor-ROI im CT und die parametrischen Karten des statischen TBR sowie der metabolischen Rate K_i zu sehen.

K_i	TBR	0.86
K_i	β	-0.33
K_i	k_1	0.54
k_1	β	-0.08

Tabelle 1. Korrelationskoeffizienten einzelner dynamischer Parameter untereinander sowie mit dem statischen TBR 3 h nach Injektion.

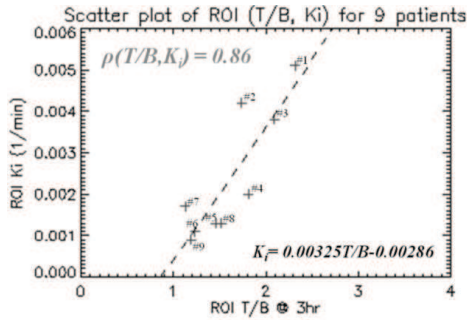


Abb. 4. Plot des mittleren K_i vs. TBR der einzelnen Patienten in der Tumorregion.

4 Schlussfolgerungen

Die Auswertungen an neun Patienten zeigen, dass die Verwendung eines TBR-Schwellwertes in Verbindung mit statischer PET Bildgebung alleine nicht in allen Fällen ausreicht um hypoxische Bereiche eindeutig zu identifizieren. Eine Kombination mit dynamischen Parametern (K_i) verspricht hier eine weitaus eindeutigere Identifikation hypoxischer Regionen.

Zwar zeigen die Ergebnisse im Mittel eine sehr deutliche Korrelation zwischen statischem TBR und dem dynamischen Parameter K_i . Gleichzeitig lassen sich aber auch an einzelnen Beispielen wie in Abb. 5 zu sehen Widersprüche zwischen dynamischer und statischer Analyse sowohl in der gesamten Tumor-ROI (Abb. 5(b)) als auch einer Sub-ROI des Tumors (Abb. 5(c)) identifizieren, in denen trotz hoher TBR niedrige K_i Werte zu beobachten sind.

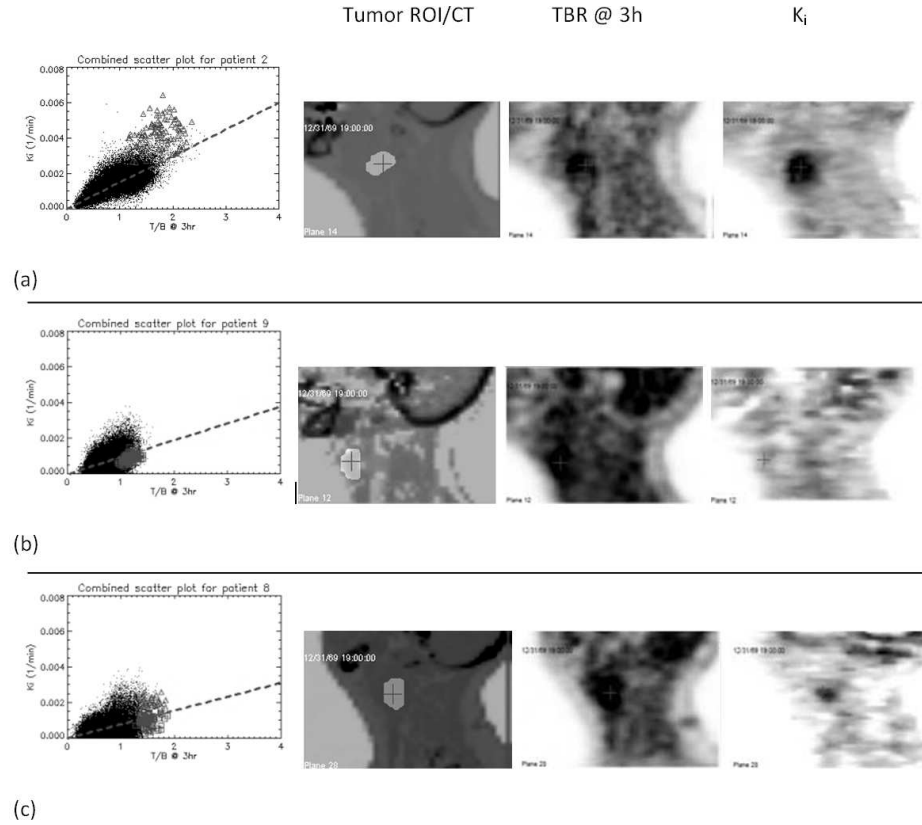
Legt man den weithin vermuteten Zusammenhang zwischen K_i und lokaler Hypoxie zugrunde, so lassen die Analysen den Schluss zu, dass die tatsächlich hypoxischen Bereiche bei hohen TBR oberhalb der individuellen Regressionsgeraden im patientenspezifischen TBR/K_i Scatterplot clustern.

Allerdings steht aufgrund der begrenzten Stichprobe von nur neun Patienten eine Auswertung auf einer deutlichen größeren Patientenpopulation aus. Daneben muss die Bestätigung des hypothetisierten Zusammenhangs zwischen K_i und tatsächlicher Hypoxie durch direkte polarografische pO_2 Messungen erfolgen.

Literaturverzeichnis

1. Nordmark M, Overgaard M, Overgaard J. Pretreatment oxygenation predicts radiation response in advanced squamous cell carcinoma of the head-and-neck. *Radiother Oncol.* 1996;41:31–9.

Abb. 5. Scatterplots dreier Patienten (K_i vs TBR) sowie die zugehörigen Schnittbilder durch die Tumor-ROI, die Segmentierung des Tumors aus dem FDG Bild überlagert dem CT, die parametrische Karte des TBR nach 3 h sowie der metabolischen Rate K_i .



2. Logan J. Graphical analysis of PET data applied to reversible and irreversible tracers. Nucl Med Biol. 2000;27:661–70.
3. Wang W, Georgi JC, Nehmeh SA, et al. Evaluation of a compartmental model for estimating tumor hypoxia via FMISO dynamic PET imaging. Phys Med Biol. 2009;54(10):3083–99.
4. Wang W, Lee NY, Georgi JC, et al. Pharmacokinetic analysis of hypoxia 18F-Fluoromisonidazole dynamic PET in head and neck cancer. J Nucl Med. 2010;51(1):37–45.

Perspective Error Correction using Registration for Blockface Volume Reconstruction of Serial Histological Sections of the Human Brain

Björn Eiben^{1,2}, Christoph Palm^{1,3}, Uwe Pietrzyk^{1,3}, Christos Davatzikos⁴
Katrin Amunts^{1,2}

¹Institute for Neuroscience and Medicine (INM-1), Research Center Jülich

²Institute for Psychiatry and Psychotherapy, RWTH Aachen

³Institute for Physics, University Wuppertal

⁴Department of Radiology, University of Pennsylvania

b.eiben@fz-juelich.de

Abstract. For accurate registration of histological sections blockface images are frequently used as three dimensional reference. However, due to the use of endocentric lenses the images suffer from perspective errors such as scaling and seemingly relative movement of planes which are located in different distances parallel to the imaging sensor. The suggested correction of those errors is based on the estimation of scaling factors derived from image registration of regions characterized by differing distances to the point of view in neighboring sections. The correction allows the generation of a consistent three dimensional blockface volume.

1 Introduction

Human post-mortem brains are sectioned with micro- and cryotomes in order to analyze the microstructure of the brain with microscopical resolution in histological and autoradiographic studies. Distortions due to sectioning and other steps of histological processing, however, are inevitable. To correct for these distortions and to perform a three dimensional (3D) reconstruction of the sections, blockface images which show the tissue block prior to cutting each section (Fig. 1(a)), proved to be an important reference modality [1, 2].

Blockface images relate directly to a 3D reference of brain tissue if the imaging setup remains constant during the whole sectioning process. In practice however, the position of the tissue block with respect to the camera is not completely fixed since the block moves against the microtome knife section to section. Furthermore the tissue plane moves away from the camera as the microtome knife is lowered iteratively.

The registration of a reference grid beneath the tissue block does not lead to full 3D consistency of the tissue as it ignores the usage of an endocentric optic and its resulting perspective errors.

2 Materials and Methods

Frozen brain tissue was sectioned into $20\ \mu\text{m}$ coronal sections with a cryostat microtome (Polycut CM 3500, Leica, Germany). The blockface images were taken by a CCD camera with a resolution of 2048×2048 pixels (Macrofire, Optronics, Goleta, USA) with an endocentric lens (APO-Componon 4.5/90, Schneider-Kreuznach, Germany). Image acquisition results in a spatial resolution in the reference-grid plane of $95\ \mu\text{m}/\text{pixel}$. The approximate distance of the lenses primary plane to the reference grid is $d = 1.23\ \text{m}$.

The knife position of the microtome is altered in z -direction from one section to the next, hence the distance of the tissue to the lenses primary plane changes equally between blockface images resulting in decreasing reproduction scales.

Utilizing homogenous coordinates to describe the imaging chain (Fig. 2, [3]), differing positions of the tissue block can be expressed by a displacement matrix \mathbf{V}_n which varies slightly for each section n (Fig. 2(a),v 2(b)). The perspective projection of a point is calculated by a multiplication with $\mathbf{P}_\mathcal{L}$. The final step is the parallel projection to the imaging sensor plane expressed by a multiplication with \mathbf{P}_Π . Thus the perspectively projected point \mathbf{p}_d can be described as

$$\mathbf{p}_d = \mathbf{P}_\Pi \cdot \mathbf{p}_c = \mathbf{P}_\Pi \cdot \mathbf{P}_\mathcal{L} \cdot \mathbf{V}_n \cdot \mathbf{p}_a = \begin{pmatrix} (x \cos(\alpha) - y \sin(\alpha) + t_x) \frac{d}{z+d} \\ (x \sin(\alpha) + y \cos(\alpha) + t_y) \frac{d}{z+d} \\ 0 \\ 1 \end{pmatrix} \quad (1)$$

$$\mathbf{P}_\Pi := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad \mathbf{P}_\mathcal{L} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{d} & 1 \end{pmatrix}; \quad \mathbf{V}_n := \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 & t_{x,n} \\ \sin \alpha & \cos \alpha & 0 & t_{y,n} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad \mathbf{p}^T := \begin{pmatrix} x^T \\ y^T \\ z^T \\ 1 \end{pmatrix}$$

As a first step the reference grid is segmented from all blockface images utilizing a variance filter followed by the analyzation of the Hough-Transformation (Fig. 1(b)). Hence the image domain Ω^I is split into grid Ω^G and background $\Omega^B = \Omega^I \setminus \Omega^G$. Thereafter the grid of each image $I_n(\mathbf{x})$, with $1 \leq n \leq N$,

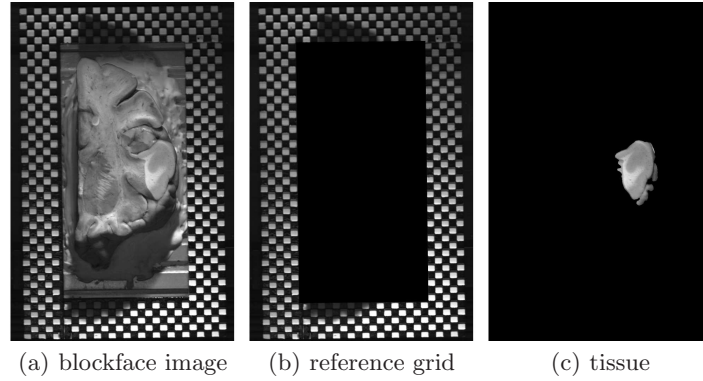


Fig. 1. Blockface images of a frozen tissue block of a human brain.

is rigidly registered to the grid of a central reference image I_c , $c \approx \frac{N}{2}$ by the transformation $\mathcal{T} : \mathbb{Z}^2 \rightarrow \mathbb{R}^2$:

$$I_n^{\mathcal{T}G}(\mathbf{x}) := I_n(\mathcal{T}_G(\mathbf{x}, \theta)) \quad \text{with} \quad \mathcal{T}_G(\mathbf{x}, \theta) := \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \mathbf{x} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (2)$$

The parameters of the rigid transformation θ are rotation by α and translation by \mathbf{t} . A masked normalized correlation metric m_{NC} guides the rigid registration process to account for illumination changes. It is calculated on those points only where the grid is present in the moving and the fixed image:

$$\Omega_n^{\mathcal{T}G} := \{ \mathbf{x} \mid (\mathbf{x} \in \Omega_c^G) \wedge (D(\mathcal{T}_G(\mathbf{x})) \in \Omega_n^G) \} \quad (3)$$

where $D : \mathbb{R}^2 \rightarrow \mathbb{Z}^2$ maps to discrete coordinate positions.

$$m_{\text{NC}}(I_n^{\mathcal{T}G}, I_c, \Omega_n^{\mathcal{T}G}) = \frac{\sum_{\mathbf{x} \in \Omega_n^{\mathcal{T}G}} (I_n^{\mathcal{T}G}(\mathbf{x}) - \bar{I}_n^{\mathcal{T}G})(I_c(\mathbf{x}) - \bar{I}_c)}{\sqrt{\sum_{\mathbf{x} \in \Omega_n^{\mathcal{T}G}} (I_n^{\mathcal{T}G}(\mathbf{x}) - \bar{I}_n^{\mathcal{T}G})^2 \sum_{\mathbf{x} \in \Omega_n^{\mathcal{T}G}} (I_c(\mathbf{x}) - \bar{I}_c)^2}} \quad (4)$$

with \bar{I}_c being the mean gray-value of I_c on the overlapping grid domain. The rigid registration parameters are optimized using an evolutionary strategy.

$$\hat{\theta}_n^G = \underset{\theta}{\operatorname{argmax}} (m_{\text{NC}}(I_n^{\mathcal{T}G}, I_c, \Omega_n^{\mathcal{T}G})) \quad (5)$$

The obtained registration results describe the block displacement in relation to the central reference image for each section. Different block positions cause different viewpoints from which the tissue block is exposed. This results in a relative movement of tissue and grid in the blockface image since they are positioned in different z -planes (compare y -positions of \mathbf{p}_b^T and \mathbf{p}_b^G with its perspectively projected points \mathbf{p}_c^T and \mathbf{p}_c^G in Fig. 2(b)). As a second step registration of the

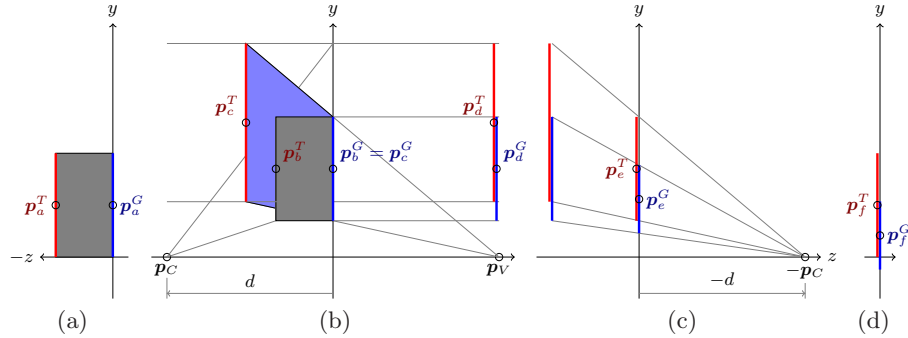


Fig. 2. Perspective distortion and correction. \mathbf{p}^T is located on the tissue plane (T) and \mathbf{p}^G on the grid (G). The block \mathbf{p}_a (a) is displaced \mathbf{p}_b , perspectively transformed \mathbf{p}_c and parallel projected \mathbf{p}_d (b). Correct scaling projects the tissue onto the corresponding grid position $\mathbf{p}_d^G = \mathbf{p}_e^T$ (c). Hence the inverted displacement of the block results in equal y -positions of the corrected point \mathbf{p}_f^T (d) and the original point \mathbf{p}_a^T .

tissue segment (Fig. 1(c)) in neighboring sections compared to the corresponding block positions allows for measuring the relative movement of grid and tissue. The metric of the registration is masked by the tissue segment

$$\Omega_{n+1}^{\mathcal{T}_T} := \{ \mathbf{x} \mid (\mathbf{x} \in \Omega_n^I) \wedge (D(\mathcal{T}_T(\mathbf{x})) \in \Omega_{n+1}^T) \} \quad (6)$$

replacing $\Omega_n^{\mathcal{T}_G}$ in (4) and resulting in the optimized parameters $\hat{\theta}_{n+1}^T$ according to (5). As consecutive sections differ in z -direction by $20 \mu\text{m}$ the assumption of no change in z holds true with an acceptable error. Finding the transformations \mathbf{A}_{n+1}^T from $\hat{\theta}_{n+1}^T$ and \mathbf{A}_n^G from $\hat{\theta}_n^G$ maps the following points to each other:

$$\begin{aligned} \mathbf{p}'_{T,n} &= \mathbf{A}_n^G \cdot \mathbf{P}_z (\mathbf{A}_n^G)^{-1} \cdot \mathbf{p}_T \\ &= \mathbf{A}_{n+1}^T \cdot \mathbf{P}_z (\mathbf{A}_{n+1}^G)^{-1} \cdot \mathbf{p}_T = \mathbf{p}'_{T,n+1} \end{aligned} \quad (7)$$

Comparing the different registrations, which on the one hand align the reference grids and on the other the tissue structures allows to calculate the scaling factor.

$$1 + \frac{z_n}{d} = \frac{t_{y,n}^G - t_{y,n+1}^G}{t_{y,n}^G - t_y^T} \quad (8)$$

Scaling each image by this estimated factor followed by the transformation detected for the reference grid corrects the perspective error (Fig. 2(c) and 2(d)) introduced in (1).

3 Results

A series of 1167 blockface images was used to evaluate the performance of the presented perspective error correction. Block displacements and relative tissue

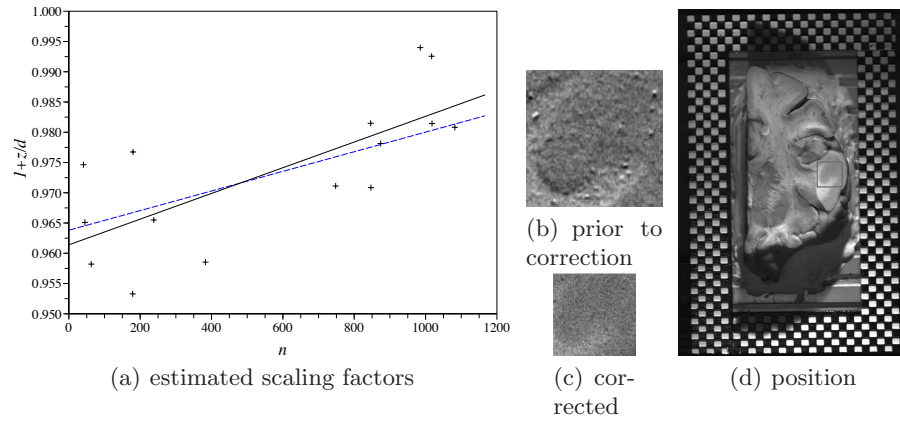


Fig. 3. Results of the depth estimation and correction. Calculations of (8) are marked with +. The linear regression is shown as a solid and the theoretical slope as a dotted line (a). (b) and (c) show the difference of the same consecutive sections without and with perspective error correction for the region marked in (d).

to grid movements were measured by mono-modal rigid image registration which was implemented using the Insight Segmentation and Registration Toolkit. A significant block displacement is necessary to measure the relative movement of tissue and grid. Therefore the evaluation of (8) was limited to sections with block displacements larger than 3 mm. With a correlation coefficient of $\rho = 0.73$ slope and axis in intercept of the scaling factor line are estimated (Fig. 3(a)). Comparing the regression with the samples, a standard deviation of $\sigma = 7.9 \cdot 10^{-3}$ results. Using knowledge about the camera distance d , section number n and the section thickness, a theoretical slope can also be calculated which supports the estimated regression line that was used for correction of the perspective error. Difference images of consecutive sections without and with perspective error correction reveal the enhancement of consistency (Fig. 3 (b-d)).

4 Discussion

The displacement of the grid plane measured by registration is – due to the perspective error – not equal to the displacement of the tissue plane located in a different distance to the point of view. The presented perspective error correction and object scaling estimation is able to calculate and correct for this error. Registration results of neighboring tissue in combination with the measured displacements allow to estimate the tissue scaling. Relative movements of tissue in neighboring sections can be effectively compensated. The perspective error gets more pronounced the thicker the block for sectioning is.

The use of telecentric instead of endocentric lenses avoids perspective errors. The usually required short distance to the object and narrow field of depth however make it not feasible for the use in a cryostat microtome. For the estimation of the scaling factor, visible tissue displacements due to perspective errors need to be distributed over the whole data set in a reasonable number. Otherwise the estimation is likely to become unreliable. The registration results need to be of high accuracy for estimating the scaling factor correctly. If this can be ensured, the presented method allows the perspective error correction without the exact knowledge of the original setting.

Acknowledgement. This study was supported by the IRTG 1328 of the Deutsche Forschungsgemeinschaft (DFG).

References

1. Amunts K, Zilles K. Atlases of the Human Brain: Tools for Functional Neuroimaging. In: Zaborsky L, Wouterlood FG, Lanciego JL, editors. Neuroanatomical Tract-Tracing 3. Springer Berlin/Heidelberg; 2006. p. 566–603.
2. Schormann T, Dabringhaus A, Zilles K. Statistics of deformations in histology and application to improved alignment with MRI. *IEEE Trans Med Imaging*. 1995;14(1):25–35.
3. Jähne B. Digitale Bildverarbeitung. Springer Berlin/Heidelberg; 2002.

Approximation des Tumormasseeffekts mittels direkt-manipulierender Free-Form Deformation

Stefan Becker, Andreas Mang, Thorsten M. Buzug

Institut für Medizintechnik, Universität zu Lübeck
becker@imt.uni-luebeck.de

Kurzfassung. In der vorliegenden Arbeit wird ein neuartiger Ansatz zur Approximation des Masseffekts von primären Hirntumoren beschrieben. Die Progression des Tumors wird mittels einer deterministischen Reaktionsdiffusionsgleichung modelliert. Die Beschreibung der raumfordernden Wirkung des Tumors erfolgt über die Kopplung der resultierenden Verteilung der Tumorzellkonzentration mit einem parametrischen Deformationsmodell, das auf der direkt-manipulierenden Free-Form Deformation basiert. Zur Generierung räumlicher Korrespondenzen wird die raumzeitliche Entwicklung der Isofläche des Tumorkernvolumens zwischen zwei Simulationsschritten mittels Landmarken verfolgt. Die Deformationen von rigiden Strukturen, wie z.B. des Schädels, wird durch starre Landmarken unterbunden. Eine Adaption der Landmarkenverteilung während des Tumorwachstums erlaubt das nachträgliche Hinzufügen neuer Landmarken. Erste qualitative Ergebnisse zeigen, dass das beschriebene Modell eine plausible Approximation der durch die Ausbreitung des Tumors hervorgerufenen raumfordernden Wirkung erlaubt.

1 Einleitung

Das Glioblastom Multiforme stellt die häufigste und aggressivste Manifestation primärer Hirntumoren beim Menschen dar. Trotz stetiger Bemühungen, Behandlungsmethoden zu verbessern, kommen heute gängige Therapien an ihre Grenzen und verbleiben in den meisten Fällen palliativ.

An dieser Stelle setzt die mathematische Modellierung der raumzeitlichen Entwicklung von Tumoren an. Ein zuverlässiges Modell würde eine Schätzung der weiteren Tumormigration erlauben, folglich Aussagen über den patientenindividuellen Verlauf zulassen und eine weitere Verbesserung der Therapieplanung ermöglichen. Die Modellierung des expansiven Verhaltens des Tumors und die Integration funktioneller Atlanten könnte weitere Aufschlüsse über mögliche Dysfunktionen während der Tumorprogression geben und somit eine Risiko-Nutzen-Analyse medizinischer Interventionen erlauben.

2 Stand der Forschung und wesentlicher Beitrag

Die Modellierung der Progression von Hirntumoren ist ein aktives Forschungsgebiet. Die mechanische Kopplung des expansiven Verhaltens des Tumors geschieht

typischerweise mittels eines lokalen Druckfeldes als parametrisierte Funktion der Zelldichte [1, 2] oder wird völlig vernachlässigt [3]. In einer aktuellen Arbeit [2] werden die Modellparameter einer vereinfachten isotropen Reaktionsdiffusionsgleichung aus patientenindividuellen, longitudinalen Bilddaten bestimmt. Das entstehende inverse Problem verwendet als Randbedingung die das Modell beschreibenden partiellen Differentialgleichungen. In [4] wird die Kopplung des expansiven Verhaltens des Tumors an die zugrunde liegende Reaktionsdiffusionsgleichung mittels einer globalen Thin-Plate-Spline-Interpolationsstrategie vorgenommen. Landmarken, die auf der Tumoroberfläche verteilt werden, und deren raumzeitliche Evolution definieren die benötigten Korrespondenzen. Die Deformation rigider Strukturen wird durch starre Landmarken unterbunden. In [5] wird der Masseffekt durch eine lokale Deformationsstrategie mittels der Free-Form Deformation durch Optimierung einer Zielfunktion approximiert. Hierbei sind Deformationen rigider Strukturen möglich.

Der wesentliche Beitrag der vorliegenden Arbeit liegt in einer Verbindung der Modelle zur Beschreibung des Tumormasseeffekts aus [4] und [5]. Dazu wird eine auf der Free-Form Deformation basierende, adaptive Erweiterung, die so genannte direkt-manipulierende Free-Form Deformation, zur Beschreibung des expansiven Verhaltens des Tumors vorgeschlagen, die eine Modellierung lokaler Deformationen erlaubt. Das in [4] vorgestellte adaptive Landmarkenschema ermöglicht eine direkte Kopplung der Deformation an die berechnete Zellkonzentration. Durch die Verwendung von starren Landmarken wird eine Deformation rigider Strukturen unterbunden.

3 Material und Methoden

3.1 Wachstumsmodell

Das verwendete mathematische Modell für die Progression des Tumors basiert auf dem in [5] vorgestellten Ansatz. Hierbei wird die raumzeitliche Veränderung der Tumorzellkonzentration auf Basis einer kontinuierlichen Reaktionsdiffusionsgleichung beschrieben. Es werden zwei wesentliche Prozesse des Tumorstwachstums modelliert, nämlich (i) die anisotrope, gewebeabhängige Diffusion maligner Zellen in das umliegende Gewebe und (ii) die Proliferation neuer Tumorzellen. Die Gewebeinformationen liefert der auf dem synthetischen MNI-Datensatz basierende Atlas [6].

3.2 Parametrisches Deformationsmodell

In diesem Abschnitt soll näher auf das verwendete Deformationsmodell eingegangen werden. Zur Beschreibung der Deformation wird dabei eine von Hsu et al. [7] vorgestellte Erweiterung der Free-Form Deformation (FFD), die so genannte direkt-manipulierende Free-Form Deformation (DMFFD), genutzt.

Sei $T: \Omega \rightarrow \mathbb{R}^3$, $\Omega \subset \mathbb{R}^3$ ein gegebenes Templatebild, welches das zu deformierende Hirn abbildet und sei $\mathbf{x} = (x_1, x_2, x_3)^T$, $\mathbf{x} \in \Omega$. Analog zur FFD wird der

Lösungsraum über ein äquidistantes Gitter $G \in \mathbb{R}^{n \times 3}$ von $n \in \mathbb{N}$ Kontrollpunkten $\{g_{k_1, k_2, k_3} \in \mathbb{R}^3 : k_i \in \mathbb{Z}, i \in \{1, 2, 3\}\}$ mit Gitterabstand $\delta \in \mathbb{R}^+$ parametrisiert. Eine Linearkombination zwischen kubischen B-Splines $\beta_p, p \in \{0, \dots, 3\}$ und den Kontrollpunkten erlaubt es, eine räumliche Abbildung $\varphi: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $\mathbf{x} \mapsto \varphi(\mathbf{x})$ zu definieren, welche die neue Position $\mathbf{y} = \varphi(\mathbf{x})$ eines Punktes $\mathbf{x} \in \Omega$ beschreibt. Hiermit ergibt sich

$$\mathbf{y} = \sum_{l_1, l_2, l_3=0}^{3,3,3} \prod_{i=1}^3 \beta_{l_i}(u_i(x_i)) g_{k_1+l_1, k_2+l_2, k_3+l_3} \quad (1)$$

wobei $u_i \in \mathbb{R}$, $u_i = x_i/\delta - [x_i/\delta]$ und $k_i \in \mathbb{Z}$, $k_i = [x_i/\delta] - 1$ In Matrixform ist (1) gegeben durch $Y = B(G + \Delta G)$ und

$$\Delta Y = B \Delta G \quad (2)$$

Hierbei bezeichnet $\Delta G \in \mathbb{R}^{n \times 3}$ die Gitterknotenverrückung und $\Delta Y \in \mathbb{R}^{m \times 3}$ die Positionsänderung der m Punkte des diskreten Bildes in die drei Raumrichtungen. Die Zeilen der Matrix $B \in \mathbb{R}^{m \times n}$ beinhalten die Werte der Basisfunktionen β_p für die Gitterknoten in Abhängigkeit von der relativen Position eines Punktes \mathbf{x} innerhalb des Gitters und bestimmen somit den Einfluss eines Gitterknotens auf die neue Position eines Punktes.

Ausgehend von der Verrückung ΔG der Gitterknoten werden somit bei der FFD Punktverrückungen ΔY erzeugt. Die DMFFD kehrt diesen Prozess um, indem einzelne Werte für ΔY vorgegeben werden. Gl. (1) zeigt, dass die Verrückung eines Punktes über die Verrückung ΔG der 64 benachbarten Gitterknoten definiert ist. Diese 64 Verrückungen müssen bei der DMFFD bestimmt werden. Sollten mehr Verrückungen von Punkten vorgegeben sein, gilt es entsprechend, ΔG für alle beeinflussten Gitterknoten zu finden. Die Lösung dieses unterbestimmten linearen Gleichungssystem ist ein Standardproblem aus der numerischen Mathematik. In der vorliegenden Arbeit wird zur Lösung ein Gradientenabstiegsverfahren verwendet [8]. Die berechnete Gitterknotenverrückung, die (2) für vorgegebene ΔY mit kleinen ΔG erfüllt, kann im Anschluss genutzt werden, um den Masseffekt zu approximieren. Sollte das Residuum der Lösung über einer Schwelle $\epsilon \in \mathbb{R}^+$ liegen, wird ΔG verworfen, da physikalisch inkorrekte Deformationen auftreten würden.

3.3 Kopplungsmodell

Wie im vorherigen Abschnitt beschrieben, werden zur Approximation des Masseffekts mittels DMFFD Punktverrückungen ΔY benötigt, die das expansive Verhalten des Tumors beschreiben. Dazu wird ein in [4] detailliert vorgestellter Ansatz verwendet. Die räumliche Verschiebung von Landmarken auf der Isofläche des Tumorkernvolumens wird genutzt (weiße Quadrate in Abb. 1 (a)), um somit korrespondierende Landmarken P^j und P^{j+1} zu den Zeitpunkten j und $j+1$ zu gewinnen. Ausgehend vom Masseschwerpunkt des Tumorkernvolumens werden diese Landmarken entlang gleichmäßig verteilter, radialer Linien

l verschoben. Die Landmarkenverrückungen zwischen zwei aufeinander folgenden Zeitpunkten entsprechen den vorzugebenden Punktverrückungen ΔY . Sie werden zur Bestimmung der Gitterknotenverrückung ΔG genutzt und erlauben somit eine Approximation des Masseffekts.

Zur Unterdrückung von Deformationen rigider Strukturen werden zusätzlich starre Landmarken eingeführt (grüne Quadrate in Abb. 1 (a)), die gleichmäßig auf rigiden Strukturen (z.B. Schädel) verteilt werden. Eine nachträgliche Verfeinerung der Landmarkenverteilung sichert eine dichte Abtastung der Tumorbefläche und eine damit einhergehende gute Approximation des Masseffekts. Schematisch ist die Landmarkendynamik zu zwei unterschiedlichen Zeitpunkten in Abb. 1 (a) dargestellt.

3.4 Modellparameter

Initial wird der Lösungsraum über ein äquidistantes Gitter mit $\delta = 10$ mm parametrisiert. Daneben werden 1733 Landmarken gleichmäßig auf den Schädel und 42 Landmarken auf der Oberfläche der Tumorkernvolumens verteilt. Die Verrückungen dieser Landmarken bestimmen die Parameter zur Berechnung der Gitterknotenverrückung. Zur Beschleunigung des Verfahrens zur Bestimmung von ΔG können einzelne starre Landmarken aus der Berechnung ausgeschlossen werden, wenn die 64 benachbarten Gitterknoten nicht durch die Verrückung von Tumorlandmarken beeinflusst werden.

4 Ergebnisse

Erste qualitative Ergebnisse des vorgestellten, parametrischen Deformationsmodells zur Approximation des Masseffekts werden in Abb. 1 dargestellt. Diese zeigt neben dem Originalbild (Abb. 1 (b)) auch die deformierten Bilder (Abb. 1(c, d)) zu zwei späteren Zeitpunkten, wobei die simulierte Tumorzellverteilung überlagert wurde. Niedrige Tumorzellkonzentrationen (blau eingefärbte Bereiche) können somit mit dem diffusen Charakter des Tumors assoziiert werden, während hohe Zellkonzentrationen (rot eingefärbte Bereiche) das Tumorkernvolumen kennzeichnen. Der Masseffekt manifestiert sich besonders auffällig

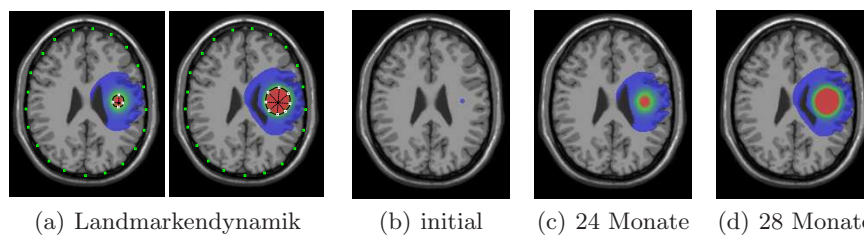


Abb. 1. Illustration der Landmarkendynamik (a) und Ergebnisse der Modellierung (c, d) unter Einblendung der Tumorzellkonzentration zu verschiedenen Zeitpunkten.

am rechten Ventrikel sowie im Bereich der Großhirnrinde. Die Anzahl der Landmarken ist zum Ende der Simulation von anfänglich 42 auf 217 Landmarken angestiegen.

5 Diskussion

In der vorliegenden Arbeit wurde ein parametrisches Deformationsmodell zur Kopplung der Reaktionsdiffusionsgleichung an die raumfordernde Wirkung des Tumors vorgestellt. Es wurde gezeigt, dass eine Beschreibung des expansiven Verhaltens des Tumors durch die Verfolgung der raumzeitlichen Evolution des Tumorkernvolumens mittels Landmarken und deren Integration in das vorgestellte Modell es generell erlaubt, den Masseffekt lokal plausibel zu approximieren. Die Verwendung von Landmarkenverrückung ermöglicht eine direkte Kopplung an das expansive Verhalten des Tumors. Die Integration von starren Landmarken erlaubt eine adaptivere Steuerung des Deformationsverhaltens von rigiden Strukturen, als es durch die Optimierung der Gitterknotenverrückung anhand der Tumorzellkonzentration [5] möglich wäre. Ein Ansatz, wie er in [9] vorgestellt wird, erlaubt voraussichtlich eine direkte Integration der Methode in ein kombiniertes Verfahren aus Tumorwachstumsmodellierung und Bildregistrierung.

Es besteht kein Zweifel daran, dass die vorgestellte Arbeit eine Machbarkeitsstudie darstellt. Eine exakte Validierung anhand von Patientendaten ist schwierig umzusetzen. Erste Ansätze für eine qualitative Evaluierung befinden sich aber im Aufbau.

Literaturverzeichnis

1. Bondiau PY, et al. Biocomputing: numerical simulation of glioblastoma growth using diffusion tensor imaging. *Phys Med Biol.* 2008;53(4):879–93.
2. Hogeia C, et al. An image-driven parameter estimation problem for a reaction-diffusion glioma growth model with mass effects. *J Math Biol.* 2008;56(6):793–825.
3. Konukoglu E, et al. Image guided personalization of reaction-diffusion type tumor growth models using modified anisotropic eikonal equations. *IEEE Trans Biomed Eng.* 2009;(accepted).
4. Becker S, et al. An adaptive landmark scheme for modeling brain deformation in diffusion-based tumor growth. In: *Proc World Congr Med Phys Biomed Eng.* vol. 25/IV; 2009. p. 41–4.
5. Becker S, et al. Tumor-Wachstumsmodellierung als parametrisches Bildregistrierproblem. In: *Proc BVM.* Heidelberg, Springer; 2009. p. 197–201.
6. Collins DL, et al. Design and construction of a realistic digital brain phantom. *IEEE Trans Biomed Eng.* 1998;17(3):463–8.
7. Hsu WM, et al. Direct manipulation of free-form deformations. In: *Proc SIGGRAPH;* 1992. p. 177–84.
8. Meister A. *Numerik linearer Gleichungssysteme. Eine Einführung in moderne Verfahren.* 3rd ed. Vieweg+Teubner Verlag; 2008.
9. Tustison NJ, et al. Directly manipulated free-form deformation image registration. *IEEE Trans Image Process.* 2009;18(3):624–35.

Bildgestützte Analyse der *in vitro*-Sedimentation agglomerierter Nanopartikel

Darius Schippritt¹, Martin Wiemann^{2,3}, Hans-Gerd Lipinski¹

¹Biomedical Imaging Group, FB Informatik, Fachhochschule Dortmund

²Physiologisches Institut, Universität Duisburg-Essen

³Institut für Biologische Emissionsbewertung (IBE), Marl

darius.schippritt@gmx.de

Kurzfassung. Es wurde eine Bildanalyse-Methode entwickelt, mit der *in vitro* Untersuchungen zur Nanopartikel-Exposition unter kontrollierten Bedingungen durchgeführt werden können. Die dafür notwendige Kenntnis über die zeitliche Verteilung der Anzahl agglomerierter Partikel lässt sich aus digitalen Mikroskopie-Aufnahmen gewinnen. Mit Hilfe einer Modellannahme können Prozess bestimmende Parameter definiert und experimentell aus den Bilddaten bestimmt werden, wobei die experimentell gewonnenen Werte dieser Parameter vollständig reproduzierbar sind. Auf diese Weise lässt sich die *in vitro* Analyse standardisieren, welches eine wichtiges Instrumentarium zum grundsätzlichen Verständnis der Wechselwirkungen von Nanostäuben und dem Immunsystem darstellt.

1 Einleitung

Sowohl tierexperimentelle Untersuchungen als auch klinische Befunde weisen darauf hin, dass Luft getragene Nanopartikel („Nanostäube“) zunehmend als erhebliches Gesundheitsrisiko einzustufen sind [1, 2]. Über die Verteilung solcher Stäube im Lungengewebe und die Auswirkung auf das Immunsystem ist bis heute jedoch nur wenig bekannt. In der Zellkultur lassen sich die toxischen Partikelwirkungen experimentell an Alveolarmakrophagen der Lunge untersuchen [3]. Eine OECD-konforme Standardisierung dieser *in vitro*-Versuche könnte dabei die Anzahl erforderlicher Tierversuche erheblich einschränken [4]. Dazu ist es aber notwendig, Makrophagen unter definierten *in vitro* Bedingungen Partikel zur Phagozytose anzubieten. Um solche standardisierten Verfahren zu etablieren, ist es unerlässlich, den charakteristischen Zeitverlauf der Partikelsedimentation unter Zellkulturbedingungen zu erfassen und mit Methoden der digitalen Bildverarbeitung zu analysieren.

2 Material und Methoden

Böhmit-Partikel (AlOOH, 180 $\mu\text{g}/\text{ml}$, Primärpartikelgröße: 40 nm) wurden in Serum freiem MEM-Medium suspendiert und in ein Zellkulturschälchen („Mes-

skammer“, 2 cm hoch) homogen verteilt gegeben. Die Sedimentation der Partikel und deren Agglomerate wurde bei 37°C, 100 % Luftfeuchtigkeit und 5 % CO₂ im Phasenkontrast (Nikon-Biostation® IM; Mikroskopfokus = Messkammerboden) bei 20facher Vergrößerung fortlaufend registriert (Rate: 1 Bild/min über 240 min) und unmittelbar nach der Bilderzeugung ausgewertet. Einzelne Nanopartikel sind so zwar nicht nachweisbar, jedoch sind die Agglomerate, die unter den experimentellen Bedingungen >99 % der Masse bilden, im Phasenkontrast-Mikroskop sichtbar (Abb. 1A). Die erzeugten Bilddaten wurden zunächst geglättet und anschließend mit Hilfe eines einfachen Schwellwert-Verfahrens so binärisiert, dass die sichtbaren Agglomerate der Nanopartikel als Objekte im Binärbild erschienen (Abb. 1B). Von den Binärobjekten wurden die Objektkanten extrahiert und zur visuellen Kontrolle den Agglomeraten im Originalbild überlagert (Abb. 1C). Anschließend wurden die Bilddaten mit dem Ziel ausgewertet, geeignete Parameter zur Erfassung der Sedimentation zu finden.

3 Ergebnisse

Ausgewertet wurde die zeitliche Verteilung der Größe $G(t)$ und der Anzahl der Agglomerate $N(t)$ auf dem Messkammerboden, wobei $N(t)$ durch einen einfachen Zählalgorithmus aus den erzeugten Binärbildern gewonnen wurde. Zunächst nahm die Anzahl der Agglomerate im Bild mit fortlaufendem Sedimentationsvorgang zu. Mit der Zeit überlagerten sich jedoch mehr und mehr Partikel am Messgefäßboden, so dass zwar die scheinbare Partikelgröße G mit der Zeit zunahm und auch die Zahl der zählbaren Partikel zunächst einem Maximum zustrebte, danach aber mit der Zeit wieder abfiel, bis eine (nahezu) konstante Partikelzahl N_0 zu beobachten war. Dieser Prozess lässt sich empirisch durch das folgende kinetische Modell sehr gut annähern

$$N(t) = \begin{cases} N_{\infty} \cdot (1 - e^{-\alpha \cdot t}) \cdot e^{-\beta \cdot t}, & \text{falls } t < t^* \\ N(t^*) = N_0, & \text{falls } t \geq t^* \end{cases} \quad (1)$$

wobei der Parameter α den Sedimentationsprozess und der Parameter β die teilweise Überlagerung der Partikel am Gefäßboden charakterisiert. Die Größe



Abb. 1. Mikroskopiebild sedimentierter, agglomerierter Nanopartikel (AlOOH); A: Originalaufnahme; B: binärisiertes Bild; C: Originalaufnahme mit überlagerter Kante.

N_∞ stellt die fiktive sedimentierte Gesamt-Partikelzahl bei vollständiger Nicht-Überlagerung (Fall $\beta = 0$) dar, während N_0 die am Boden der Messkammer tatsächlich zählbare Partikelzahl am Ende des Sedimentationsprozesses ist. Die Abb. 2 zeigt exemplarisch den zeitlichen Verlauf von $N(t)$ (Modell und tatsächlich gezählte Partikel) für ein Experiment mit agglomerierten nanoskalierten ALOOH-Partikeln. Mit Hilfe üblicher Least-Square-Approximationsverfahren wurden die Parameter der Modellfunktion (N_0 , N_∞ , α und β) an die Messdaten angepasst. Besonders wichtig ist dabei die Kenntnis darüber, zu welchem Zeitpunkt t^* der Wert N_0 erreicht ist (Ende des Sedimentationsprozesses). Er kann aus der zeitlichen Verteilungsfunktion $N(t)$ ermittelt werden.

Aus der Zeitreihenanalyse von $N(t > t^*)$ kann man zudem auf die Einflüsse von Konvektionen am Boden des Gefäßes schließen, weil sich durch Konvektionsbewegungen die eigentlich konstante Gesamtpartikelzahl N_0 scheinbar zyklisch geringfügig mit der Zeit ändert (Partikel wandern zeitweise aus der Fokusebene des Mikroskops; das Phänomen tritt übrigens während des gesamten Messvorgangs auf). Diese zeitlichen Variationen von N_0 lassen sich gut durch eine Normalverteilung $N(N_0, \sigma_0^2)$ approximieren, deren Varianz ein Maß für die Konvektion am Gefäßboden ist. Schließlich lässt sich mit Hilfe der Zeit t^* auch die mittlere Sedimentationsgeschwindigkeit \bar{v} schätzen. Sie ist definiert als Quo-

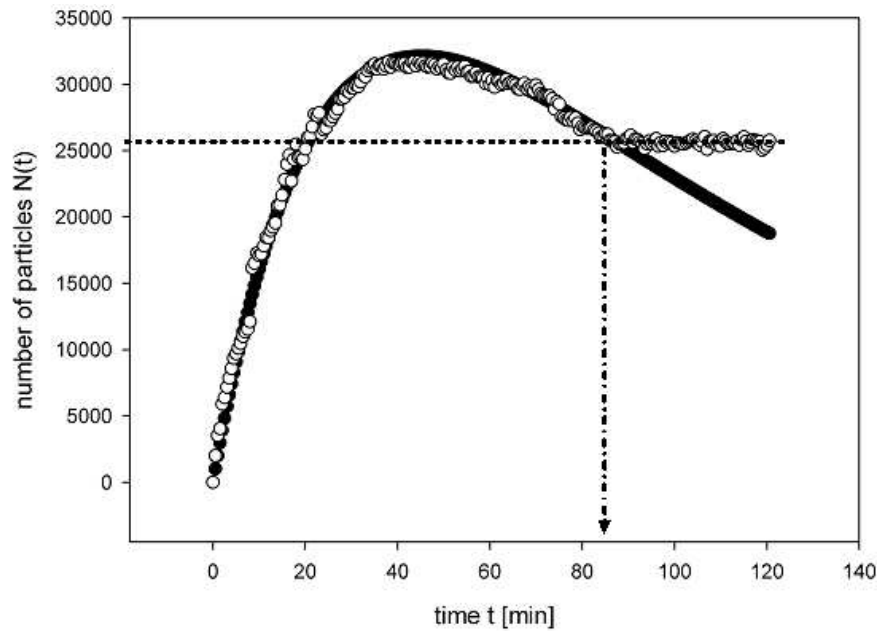


Abb. 2. Zahl agglomerierter Nanopartikel N in Abhängigkeit von der Zeit t .
 Modell: ●; korrespondierende gemessene Zellzahlen: ○;
 Parameter: $t^* = 85$ min (Pfeil); $N_0 = 25850$; $\alpha = 0,0273 \text{ min}^{-1}$; $\beta = 0,0112 \text{ min}^{-1}$;
 $N_\infty = 75267$; $\bar{v} = 235 \mu\text{m}/\text{min}$; $\sigma_{N_0}^2 = 362$.

tient von Messkammerhöhe h und Zeit t^* . Da h eine Apparatekonstante ist, wird \bar{v} letztlich nur durch t^* bestimmt.

Die Überprüfung der Modellannahmen erfolgte unter gleichbleibenden experimentellen Bedingungen (z.B. Nanopartikelsorte, mittlere Größe der Agglomerate, Temperatur), wobei alle Experimente mehrfach wiederholt durchgeführt und die Prozess-Parameter nach den beschriebenen Verfahren ermittelt wurden. Dabei zeigte sich, dass die gewonnenen Parameterwerte N_0 , N_∞ , α , β , \bar{v} bzw. t^* normal-verteilt waren, wobei die maximale Abweichung vom jeweiligen Parameter-Mittelwert höchstens 5% betrug.

4 Diskussion

Für *in vitro* Untersuchungen zur Nanopartikelexposition unter kontrollierten Bedingungen konnten mit Methoden der digitalen Bildverarbeitung eine Reihe von kinetischen Parametern identifiziert werden, die sich mit Hilfe eines geeigneten Modells aus der gemessenen Anzahl sedimentierter Partikel reproduzierbar bestimmen lassen. Empirisch haben sich insbesondere vier Parameter als wichtig erwiesen: die Gesamtsedimentationszeit t^* , die daraus abgeleitete mittlere Sedimentationsgeschwindigkeit \bar{v} , die final am Messgefäßboden verfügbare Partikelzahl N_0 sowie die fiktive Zahl N_∞ . Diese Parameter erlauben es, eine standardisierbare experimentelle Ausgangsbasis zu schaffen, so dass Wechselwirkungen dieser Nanopartikel mit vitalen Makrophagen unter kontrollierten *in vitro*-Bedingungen untersucht werden können.

Danksagung. Das Forschungsprojekt wurde mit Mitteln des Ministeriums für Innovation, Wissenschaft, Forschung und Technologie des Landes NRW gefördert.

Literaturverzeichnis

1. Xia T, Li N, Nel A. Potential health impact of nanoparticles. *Annu Rev Public Health*. 2009;30:137–50.
2. Song Y, Li X, Du X. Exposure to nanoparticles is related to pleural effusion, pulmonary fibrosis and granuloma. *Eur Respir J*. 2009;34:559–67.
3. Bruch J, Rehn S, Rehn B, et al. Variation of biological responses to different respirable quartz flours determined by a vector model. *Int J Hyg Environ Health*. 2004;207:203–16.
4. Hartung T. Toxicology for the twenty-first century. *Nature*. 2009;460:208–12.

Orts-relationale SIFT-Hierarchien zur Ähnlichkeitsbestimmung mit Graph-Matching

Armin Fritsche, Benedikt Fischer, Thomas M. Deserno

Institut für Medizinische Informatik, RWTH Aachen
afritsche@imib.rwth-aachen.de

Kurzfassung. Zur Bestimmung von Bildähnlichkeiten werden häufig charakteristische Punkte berechnet. Ein bewährtes Verfahren zur Extraktion solcher Punkte ist die Scale Invariant Feature Transform (SIFT). Gerade auf medizinischen Grauwertbildern liefert dieses Verfahren aber unbefriedigende Ergebnisse. In dieser Arbeit wird ein neuer Ansatz vorgestellt, mit dem auch strukturelle Informationen über die Verteilung der SIFT-Punkte im Bild gespeichert und verglichen werden können. Das Verfahren wurde zur Knochenaltersschätzung auf Handradiographien angewendet. Im Vergleich zum Standard SIFT wurde Recall um 4 % verbessert.

1 Einleitung

Zum Vergleich zweier Bilder existieren zahlreiche Ansätze. Die meisten Methoden basieren dabei auf der Segmentierung der Bilder und dem Vergleich der Regionen. Die Beschreibungen der Regionen sind aber häufig anfällig gegen Rotationen, Skalierungen oder Formänderungen. Daher werden anstelle der Regionen auch charakteristische Punkte, sog. Interesting Points (IPs), extrahiert. Eine weit verbreitete Technik zur Bestimmung und Beschreibung dieser Punkte ist das SIFT-Verfahren [1]. Die Deskriptoren dieser Punkte sind größtenteils invariant gegenüber Rotation, Skalierung und kleineren Änderungen des Betrachtungswinkels. Jeder SIFT-Punkt wird dabei durch seinen Deskriptor, seine Orientierung und seine Koordinaten im Skalenraum (x, y, σ) beschrieben, wobei σ seine Skalierung ist. Der Deskriptor ist ein reellwertiger Vektor. Zwei Bilder werden verglichen, indem man die SIFT-Punkte miteinander vergleicht. Der Vergleich wird dabei über die euklidische Distanz der Deskriptorvektoren vorgenommen. Ein Matching zwischen zwei Nearest Neighbour Punkten wird vorgenommen, wenn ihre Distanz unter einem Schwellwert liegt.

Im medizinischen Bereich ist dieses Verfahren zur Registrierung von MRT-Aufnahmen des Gehirns getestet [2]. Die Registrierung fand dabei allerdings nur auf künstlich deformierten Bildern statt. In [3] wird eine Klassifikation von medizinischen Bildern anhand von SIFT-Features vorgenommen. Trotz akzeptablen Ergebnis wird empfohlen, andere Deskriptoren, wie etwa Thumbnails der Bilder, zu benutzen. Ferner wird eine Aufteilung in grobe Bildklassen (z.B. Handradiographie, Schädel-CT) vorgenommen [4]. Hierbei werden einerseits künstlich

deformierte Fingerepiphysen miteinander verglichen und andererseits wird das Knochenalter durch einen Vergleich mit Epiphysen bekannten Alters bestimmt. Während im ersten Schritt noch ca. 87% der SIFT-Punkte korrekt zugeordnet werden konnten, wird bei der Altersbestimmung nur bei 2% der Bilder eine korrekte Schätzung erreicht.

Bei bisherigen Arbeiten in der medizinischen Bildverarbeitung wurde das SIFT-Verfahren also entweder nur zur groben Klassifikation oder auf künstlich veränderten Bildern erfolgreich angewendet. Für die meisten klinischen Anwendungen scheint das Verfahren nicht genügend Informationen bereitzustellen. Einer der Gründe ist die lokale Betrachtungsweise des Verfahrens, bei der die Verteilung der Punkte im Bild außer Acht gelassen wird. Diese Arbeit beschäftigt sich daher mit einer Methode, um Informationen über die Lage der IPs im Bild in einen Bildvergleich mit einbeziehen zu können.

2 Material und Methoden

In [5] werden Bilder über Bildregionen verglichen. Dabei werden nicht nur einzelne Bildregionen betrachtet, sondern auch Beziehungen zwischen den Regionen. Alle Regioneninformationen werden in einem gewichteten Graphen gespeichert. Der Vergleich wird über ein Graphmatching vorgenommen. Das selbe Prinzip wird hier verwendet, indem statt Informationen zu den Regionen und ihrer Inklusionsrelation eine entsprechende Baumstruktur der SIFT-Punkt-Hierarchie erzeugt wird, so dass der Bildvergleich durch Graphmatching erfolgen kann.

2.1 Grapherzeugung

Zunächst werden auf den Anfragebildern die SIFT-Punkte berechnet. Für jeden dieser Punkte wird dann ein Knoten im Graphen erzeugt. Die Knoten werden mit einem Deskriptor, seiner Orientierung und seinen Koordinaten im Skalerraum annotiert. Die Kantenstruktur wird aus der adaptiven nicht-maximalen Suppression (ANMS) [6] abgeleitet. Dort dient die ANMS einerseits zur Unterdrückung schlechter IPs und andererseits zur besseren Verteilung der IPs im Bild. Das Prinzip ist hierbei, dass IPs, die in ihrer lokalen Nachbarschaft nicht maximale Skalierung haben, aussortiert werden. Während bei Brown die Aussortierung über Löschung abläuft, wird in dieser Arbeit eine hierarchische Struktur eingeführt. Dafür wird um jeden SIFT-Punkt ein Kreis mit Radius r gelegt, der seine lokale Nachbarschaft definiert. Die Kreise um die SIFT-Punkte wachsen im Laufe der Grapherzeugung sukzessiv an. Ist ein Kreis soweit gewachsen, dass er nicht nur einen, sondern mehrere SIFT-Punkte enthält, wird der SIFT-Punkt mit maximalem σ -Wert gesucht und hierarchisch über alle anderen SIFT-Punkte des Kreises erhöht. Das heißt, im Graphen wird der Knoten des maximalen SIFT-Punktes zum Vaterknoten aller anderen Knoten in der selben Nachbarschaft. Alle Knoten bzw. SIFT-Punkte, die in ihrer Nachbarschaft nicht mehr maximal sind, entfallen aus der Betrachtung (Abb. 1).

Da es durch das SIFT-Verfahren zu mehreren IPs mit den selben Koordinaten, aber mit verschiedenen Orientierungen kommen kann, muss auch dieses in der Baumstruktur repräsentiert werden. Solche Knoten werden mit einem neuen Kantentyp, den Adjazenzkanten verbunden. Diese Punkte können beim Bildvergleich beliebig vertauscht werden. Um stets eine eindeutige Baumstruktur zu erhalten (abgesehen von den Adjazenzkanten), muss abschließend ein künstlicher Wurzel-Knoten eingefügt werden. Dies wird etwa dann nötig, wenn es mehrere maximale SIFT-Punkte im Bild gibt.

2.2 Relationale Merkmale und Bildvergleich

Durch die Struktur des Graphen sind schon erste Informationen über die Verteilung der IPs im Bild gespeichert. Diese geben aber zu wenig Aufschluss über die Beziehungen der Punkte untereinander. Daher werden beim Vergleich zweier Graphen zusätzlich relationale Merkmale berechnet und als Kantengewichtung in den Graphen aufgenommen. Relationale Merkmale beschreiben die Beziehung zweier IPs zueinander. Hierbei ist darauf zu achten, dass alle verwendeten Merkmale aus den Knotengewichtungen berechnet werden können und jeweils in Relation zu den SIFT-Merkmalen stehen. Durch Wahl anderer Relationen (z.B. Farbvergleich, Ortsvergleich) würde man riskieren, die Invarianzeigenschaften der SIFT-Punkte zu verlieren. Für diese Arbeit wird die Orientierungsdifferenz, die Skalierungsdifferenz, die Entfernung in Relation zu ihrer Skalierung, die Lage

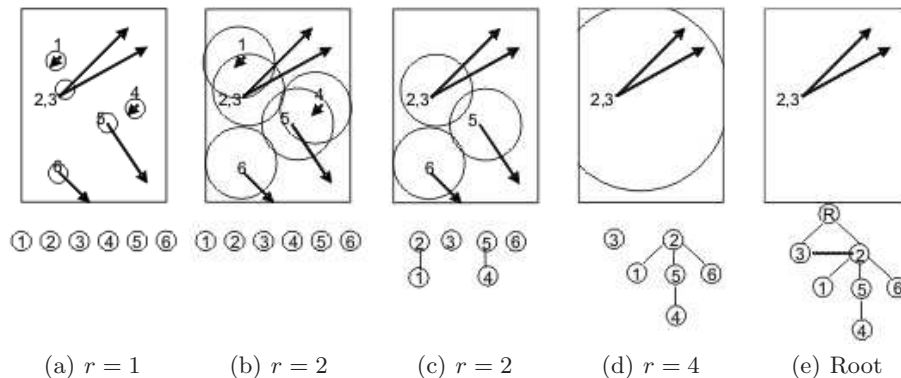


Abb. 1. (a) Jeder SIFT-Punkt wird durch seine Orientierung und Skalierung dargestellt. Um jeden Punkt wird ein Kreis gelegt. (b) Die Kreise wachsen, bis mehrere SIFT-Punkte verschiedener Skalierung in einem Kreis liegen. (c) Der jeweils maximale SIFT-Punkt in einem Kreis erhöht sich in der Hierarchie. Ist der maximale Punkt nicht eindeutig, wird einer zufällig gewählt. Punkte, die in ihrer Nachbarschaft mit Radius r nicht mehr maximal sind, fallen aus der Betrachtung heraus. (d) Dieses Verfahren wird immer weiter fortgesetzt, bis nur noch maximale Punkte im Bild existieren. (e) Abschließend werden noch die adjazenten Knoten 2 und 3 mit Adjazenzkanten verbunden, und es wird ein Wurzel-Knoten eingefügt, der Vater der Knoten 2 und 3 ist.

in Relation zur Orientierung und die euklidische Distanz zwischen den Deskriptoren verwendet.

Für zwei Bilder bekommt man demnach zwei knoten- und kantengewichtete Graphen, die sich mit einem Matchingverfahren vergleichen lassen. Hierbei ist zu beachten, dass Knoten, die über Adjazenzkanten verbunden sind, beliebig ausgetauscht werden können. In dieser Arbeit wurde ein Graphmatchingverfahren angewendet, welches auf evolutionärer Spieltheorie basiert [5]. Die Ausgabe des Graphmatchers sind schließlich Knoten- und somit auch SIFT-Punktzuordnungen.

Diese Punktzuordnungen werden aus Epiphysen-ROI von Handradiographien ausgewertet (Abb. 2). Auf Grundlage der Epiphysenvergleiche zu Epiphysen mit bekannten Alter lässt sich schließlich eine Schätzung des Knochenalters von Epiphysen unbekanntes Alters vornehmen. Die Testmenge besteht aus 98 Epiphysen unterschiedlichen Alters. Hierbei wird jede ROI mit jeder anderen verglichen. Zur Evaluation werden einerseits die Anzahl der gefundenen Matchings gezählt, andererseits werden korrekte Zuordnungen ermittelt und in Relation zu den Gesamtzuordnungen (Precision) und zu der Anzahl der extrahierten IPs (Recall) gesetzt. Zur Bewertung wird das Ergebnis mit dem Standard SIFT-Verfahren [1] verglichen.

3 Ergebnisse

Die Ergebnisse aus Tabelle 1 basieren auf den insgesamt 9506 Bildvergleichen. Die Bewertung, ob ein Matching korrekt ist, wurde in beiden Fällen anhand einer Ground Truth vorgenommen (Abb. 2). Während bei der normalen SIFT-Methode eine leicht bessere Precision erreicht wird, lässt sich der Recall beim Graphvergleich annähernd verdoppeln. Im Einzelnen wurden im Standard SIFT-Verfahren ca. 2,03 Punktmatchings pro Bildvergleich gefunden (0,75 korrekte). Im Graphmatchingansatz werden durchschnittlich 5,16 Punktmatchings pro Bildvergleich gefunden (1,53 korrekte).

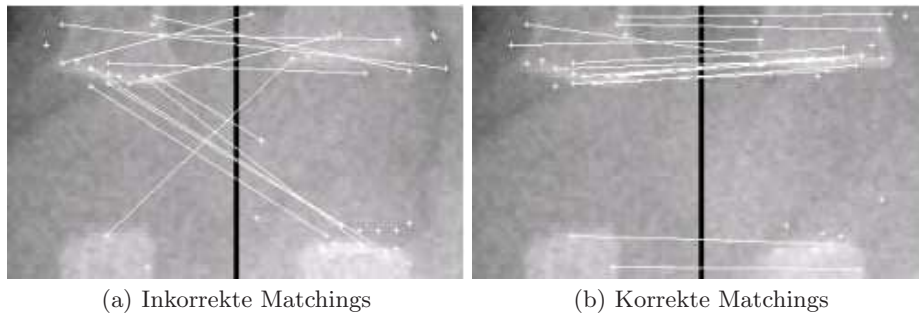


Abb. 2. Beispiele für zwei Epiphysen-ROI Vergleiche mittels SIFT-Graphmatching. Zur besseren Ansicht wurde der Kontrast nachträglich verstärkt.

Tabelle 1. Ergebnisse.

Verfahren	Matchings		Precision	Recall
	gefunden	richtig	in %	in %
Standard-SIFT	19314	7205	23,53	5,20
SIFT-Graphen	49089	14583	22,20	9,44

4 Diskussion

Die schlechte Precision der verglichenen Verfahren kann durch den Aufbau der SIFT-Deskriptoren erklärt werden. Diese bestehen aus Gradientenhistogrammen. Sind sich mehrere Gradienten sehr ähnlich, etwa an den Knochenkanten, so kann keine eindeutige Zuordnung gewählt werden. Andererseits werden durch verschiedene Intensitätsänderungen auch viele Matchings fälschlich verworfen. Durch die Betrachtung der Lage der Punkte, werden einerseits ähnliche Punkte unterscheidbar, andererseits wird die harte Bedingung der korrekten Gradienten aufgeweicht. Daher können mehr Zuordnungen berechnet werden.

Insgesamt konnte ein Verfahren gezeigt werden, welches IPs um weitere Bildinformationen ergänzt und es daher auf neuen Anwendungsgebieten einsetzbar macht. Für weitere Arbeiten empfiehlt sich eine genauere Untersuchung der relationalen Merkmale, da hier zunächst recht einfache gewählt wurden, sowie ein Wechsel der IPs.

Danksagung. Diese Arbeit entstand im Rahmen des DFG Projektes „Strukturierte Prototypen in Radiologischer Routine“ (Le 1108/9)

Literaturverzeichnis

1. Lowe DG. Distinctive image features from scale-invariant key points. *Int J Computer Vis.* 2004;60(2):91–110.
2. Moradi M, Abolmaesumi P, Mousavi P. Deformable registration using scale-space keypoints. *Proc SPIE.* 2006;6144:791–8.
3. Deselaers T, Keysers D, Ney H. Features for image retrieval: An experimental comparison. *Inf Retr Boston.* 2008;11(2):77–107.
4. Wittenhagen M. Verwendung von SIFT-Features für die Evaluierung von Bildähnlichkeiten [Studienarbeit]. Institut für Medizinische Informatik, RWTH Aachen; 2008.
5. Fischer B, Fritsche A, Thies C, et al. Evolutionäres Graphmatching zur Handknochen-Identifikation. *Proc BVM.* 2009; p. 26–30.
6. Brown M, Szeliski R, Winder S. Multi-image matching using multi-scale oriented patches. *Proc CVPR.* 2005;1:510–7.

Thallium-Stress, Technetium-Rest Protokoll für Cardiac SPECT

Phantommessungen und erste Patientendaten für ein Simultaneous-Dual-Isotope-Protokoll

T. Dey^{1,2}, H. Wiecek², B.E. Backus³, L. Romijn³, R. Bippus²,
J.F. Verzijlbergen³, T. Aach¹

¹Lehrstuhl für Bildverarbeitung, RWTH Aachen University, Germany

²Philips Technology Research Laboratories, Aachen, Germany

³St. Antonius Hospital Nieuwegein, Nieuwegein, The Netherlands

thomas.dey@philips.com

Kurzfassung. Bei dem nuklearmedizinischen funktionalen Bildgebungsverfahren SPECT werden zur Diagnostik der koronaren Herzkrankheit (KHK) üblicherweise zwei getrennte Aufnahmen in Ruhe und unter Belastung verglichen. In diesem Beitrag wird anhand von Phantommessungen und ersten Patientendaten ein Protokoll untersucht, bei dem mit nur einer SPECT-Aufnahme der Ruhe- und Belastungszustand gleichzeitig extrahiert werden können. Dies wird durch Applikation der beiden Tracer Tc-99m in Ruhe und Tl-201 unter Belastung sowie durch ein Rekonstruktionsverfahren mit einer auf Monte-Carlo Simulationen beruhenden Streukorrektur ermöglicht. Dadurch kann sowohl die für eine Studie benötigte Zeit reduziert, wie auch die Registrierung der Bilder der beiden Aufnahmen umgangen werden. Die hier vorgestellten Phantom- und Patientendaten zeigen eine gute für die klinische Beurteilung ausreichende Bildqualität und geben Anlass für weitere klinische Studien.

1 Einleitung

SPECT ist das weltweit meist verwendete funktionale Bildgebungsverfahren zur Diagnostik der KHK. Bei diesem Verfahren werden dem Patienten Radioisotope injiziert, welche sich abhängig von der Durchblutung des jeweiligen Herzmuskelbereiches anreichern. So können minderperfundierte Läsionen erkannt werden. In der klinischen Praxis werden in der Regel zwei SPECT-Aufnahmen, eine unter Belastung (Stress) und eine weitere im Ruhezustand (Rest), durchgeführt. Aus dem Vergleich beider Aufnahmen können Narbengewebe oder Ischämien unterschieden werden. Beide Untersuchungen werden nacheinander unter Gabe von verschiedenen Dosen von Radiopharmaka (z.B. Tc-99m Sestamibi oder Thallium-201) in Ein- oder Zweitagesprotokollen durchgeführt [1]. Aufgrund der höheren Emissionsenergie von Tc-99m (140 keV) im Vergleich zu Tl-201 (Hauptemissionslinien bei etwa 72 keV, Nebenemissionslinie bei 167,4 keV) werden die von Tc-99m emittierten γ -Quanten weniger stark im Patienten gestreut. Darüber

hinaus können aufgrund der kürzeren Halbwertszeit höhere Dosen an Tc-99m bei gleicher Strahlenbelastung appliziert werden. Daher liefern Tc-99m im Vergleich zu Tl-201-Aufnahmen eine bessere Bildqualität. In [2] wird ein iteratives Ordered Subset Expectation Maximization (OSEM)-Rekonstruktionsverfahren mit einer auf Monte-Carlo-Simulationen und effektiven Scatter-Quellen basierenden Streukorrektur vorgestellt, welches auch Schwächung im Patienten und das Blurring durch die Kollimator-PSF (Point-Spread Funktion) korrigiert. Mit diesem Verfahren ist auch die gleichzeitige Rekonstruktion (SDI: Simultaneous Dual Isotope) der Verteilungen von Tc-99m und Tl-201 im Patienten möglich. So entfällt die Registrierung der Ruhe- und Belastungsakquisitionen sowie die damit verbundene Problematik. In [3] wird ein Protokoll (Standard-SDI) mit Tl-Injektion in Ruhe und Tc-99m unter Belastung beschrieben. Bei diesem Protokoll führt das hohe Verhältnis von Tc-99m zu Tl-201 jedoch aufgrund der Streuung von Tc emittierten Photonen in die Tl-Energiefenster zu einer schlechten Bildqualität der Tl-201 Bilder. In diesem Beitrag wird deshalb die Bildqualität eines alternativen SDI Protokolls (Reverse-SDI) mit Tc-99m Injektion im Ruhezustand und Tl-201 unter Belastung anhand von Phantomdaten und ersten Patientendaten untersucht. Abb. 1 zeigt den zeitlichen Ablauf und die verwendeten Dosen für beide Protokolle. Im Standard-SDI Protokoll wird unter realistischen Annahmen für die Aufnahme im Myocardium, ein Tc:Tl-Verhältnis von etwa 9:1 erwartet. Für das Reverse-SDI Protokoll ergibt sich entsprechend ein Verhältnis von etwa 0,6:1, wenn man die höhere Aufnahme von Tl im Myocardium unter Belastung berücksichtigt. Eine Verkleinerung des Tc:Tl Verhältnisses ließe sich im Standard-SDI Protokoll nur durch eine Reduktion der Tc Dosis erreichen. Wesentlich höhere Tl-Dosen sind aufgrund der höheren Strahlenbelastung durch Tl nicht möglich.

2 Material und Methoden

Mit dem Experiment soll der Einfluss des Tc:Tl-Verhältnisses im Patienten auf die Bildqualität untersucht und geprüft werden, ob bei dem für das Reverse-SDI erwartete Tracerverhältnis eine für die klinische Beurteilung geeignete Bildqualität erreicht werden kann. Dazu wird ein Torsophantom mit einem Herzeinsatz verwendet. Dieses enthält neben einer Kammer zur Simulation der Leber noch zwei mit Styropor gefüllte Lungenkammern und einen Teflon[®]-Zylinder zur Nachbildung der Abschwächung durch das Rückgrat. Der Herzeinsatz selbst besteht aus einer inneren Kammer für das linke Ventrikel und einer äußeren Myocardkammer. In diese Kammer werden zwei Läsionen in Form befüllbarer

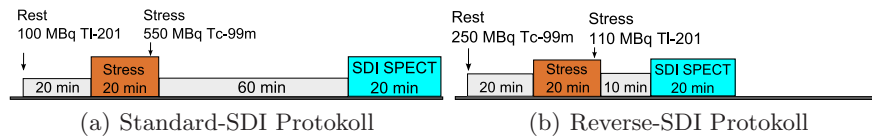


Abb. 1. Ablauf der beiden SDI Protokolle.

Tabelle 1. Läsionen im Herzeinsatz des Torsophantoms.

Läsion	Abmessung	befüllbares Volumen	Lage
Läsion 1	2 cm Höhe, 45°	5,6 ml	anterior, nahe Apex
Läsion 2	2 cm Höhe, 90°	10 ml	inferior, nahe Basis

Tabelle 2. Konzentrationen in kBq/ml der Isotope zu verschiedenen Messzeiten.

	1.Messung		2.Messung		3.Messung	
	Tc	Tl	Tc	Tl	Tc	Tl
Myocardium	240,0	111,9	65,9	100,6	52,5	98,7
Läsionen	240,0	0	65,9	0	52,5	0
Leber	56,8	18,0	15,6	16,2	12,4	15,9
Hintergrund	17,0	5,2	4,7	4,7	3,7	4,6
Lungen	6,8	2,0	1,9	1,9	1,5	1,8

Plastikbehälter eingebracht. Die Ausdehnung und Lage der Läsionen sind in Tab. 1 angegeben. Die Gradzahl gibt den Öffnungswinkel des Kreissectors bei horizontalem Schnitt durch den Herzeinsatz an. Die Läsionen enthalten kein Tl, um Ischämien zu simulieren. Durch die unterschiedlichen Halbwertszeiten (Tc-99m: 6h und Tl-201: 72h) können verschiedene Tracerverhältnisse in Abhängigkeit des Messzeitpunktes erreicht werden. Tab. 2 fasst die Aktivitäten zu den verschiedenen Aufnahmezeiten zusammen. Die zweite und dritte Messung bilden das Tc:Tl-Verhältnis des Reverse-SDI Protokolls (0,66:1 bzw. 0,53:1 im Myocardium) ab, während die erste ein Verhältnis von Tc:Tl von 2,1:1 im Myocardium, ähnlich dem des Standard-SDI Protokolls, aufweist. Die SPECT-Aufnahmen werden mit einer Philips CardioMD-Kamera (LEHR-Kollimator, 64 Frames, 20 Sekunden pro Frame) durchgeführt. Zur Rekonstruktion werden die Energiefenster EM: 126–153 keV (Tc-99m Fenster), W1: 156,3–181,68 keV (oberes Tl-201 Fenster) und W2: 64,8–79,2 keV (unteres Tl-201 Fenster) aufgenommen. Für die Schwächungskorrektur wird, zeitgleich zur SPECT-Aufnahme, die Abschwächung einer verfahrenbaren Gd-Linienquelle gemessen (Philips Vantage System). Die Rekonstruktion der Schwächungskarte aus den Transmissionsfenstern wird mit der AutoSpect+ Software der Philips Jetstream Workstation (JSWS) durchgeführt. Für die Rekonstruktion der Tc und Tl Verteilungen wird der in [2] beschriebene Rekonstruktions-Algorithmus verwendet. Zur klinischen Beurteilung wird die 5. Iteration (bei 8 Subsets und 64 Projektionen) verwendet, auf welcher noch Butterworth-Filterung (Grenzfrequenz 0,5 cycles/pitch) vorgenommen wird. Um eine in der Klinik übliche Darstellung zu erhalten, werden die Bilder mit dem AutoQuant-Paket der JSWS reorientiert und in Form von Kurz- und Längsachsenschnitten sowie zusätzlich als Polarplot ausgegeben. Die Begutachtung dieser Bilder erfolgt durch zwei erfahrene Mediziner. Zusätzlich werden die Raten der Tracer im Myocardium des Phantoms und im Patienten bestimmt. Dazu wird zunächst eine Kalibrierung mit Spritzen als Punktquellen

in einem Streuphantom (Nylonzylinder, Durchmesser 20 cm), jeweils gefüllt mit Tl und Tc, durchgeführt. Die Segmentierung des Myocardiums erfolgt auf den reorientierten Bildern mittels eines einfachen Schwellwertverfahrens, welche die Tl-freien Läsionen ausschließt.

3 Ergebnisse

In Abb. 2 sind die Ergebnisse der Schnittbilder für die Phantomdaten (Abb. 2(a-c)) und ein Patientenbeispiel (Abb.2(d)) dargestellt. Dabei zeigt Abb. 2(a) das Ergebnis der ersten und Abb. 2(b,c) das der zweiten und dritten Aufnahme. In den Abb. sind nur ausgewählte Schnitte gezeigt, in denen die Läsionen im Phantom sichtbar sind. In allen Bildern zeigt die obere Reihe jeweils die Belastungsaufnahme mit Tl und die untere die Ruheaufnahme mit Tc. Durch die Phantombilder wird die theoretisch erreichbare Bildqualität veranschaulicht. In Tab. 3 sind die gemessenen Verhältnisse von Tc:Tl für Phantom- und erste Patientenmessungen zusammengefasst.

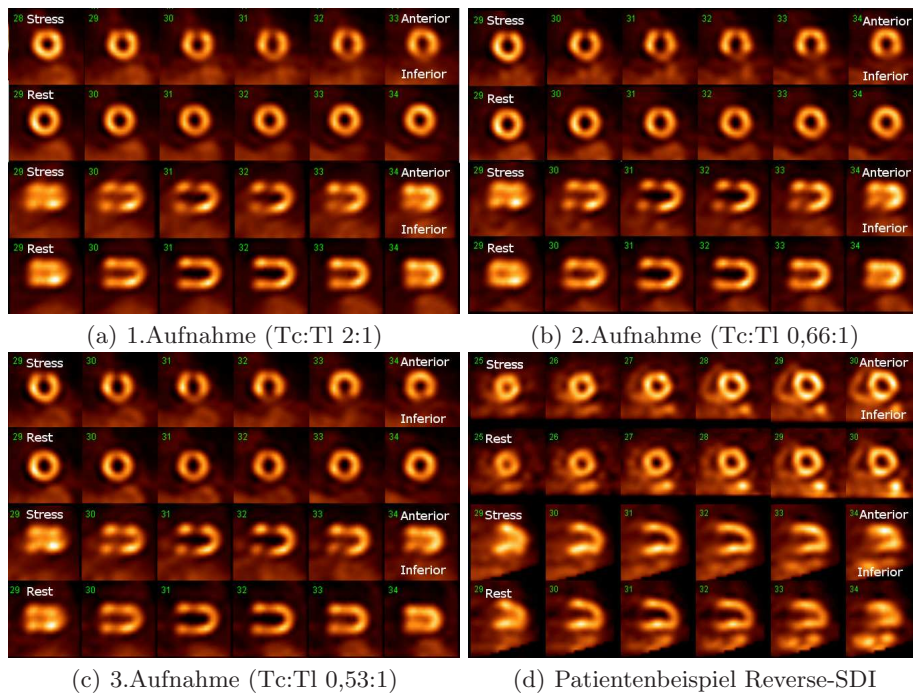


Abb. 2. Schnittbilder aus Phantom- und Patientenaufnahmen.

Tabelle 3. Verhältnisse von Tc:Tl im Myocardium.

	erwartet gemessen	
1. Phantom	2,14	2,13
2. Phantom	0,66	0,68
3. Phantom	0,53	0,56
5 Patienten	$\approx 0,6$	0,49–0,58

4 Diskussion

In Abb. 2(a) ist die inferior gelegene Läsion nur schlecht zu erkennen. In den beiden anderen Aufnahmen mit geringerem Tc:Tl Verhältnis ist dies problemlos möglich. Insgesamt ist die Qualität der Belastungsbilder in der zweiten und dritten Aufnahme deutlich besser (Abb. 2(b,c)). Zwischen diesen beiden sind kaum Unterschiede festzustellen. Die Ergebnisse zeigen, dass die Bildqualität für eine klinische Beurteilung der Aufnahmen völlig ausreichend ist. Von klinischen Experten wurde die Phantomaufnahme mit einem Tc:Tl-Verhältnis von 0,66:1 am besten bewertet. Auch die ersten Patientenbilder (Abb. 2(d)) sind vielversprechend und mit den Phantomdaten vergleichbar. Durch die zwei- bis dreimal höhere Aufnahme von Tl unter Belastung kann trotz der kleinen Tl-Dosis von 110 MBq eine genügend hohe Zählrate für eine gute Thalliumbildqualität erreicht werden. Von Tc wird keine Redistribution durch die nach der Applikation durchgeführte Belastung erwartet. Dies soll durch weitere klinische Studien validiert werden. Ohne Redistribution bleibt die Verteilung im Ruhezustand auch nach der Belastungsphase für die später durchgeführte Messung erhalten. Ein weiterer Vorteil des Reverse-SDI Protokolls ist die Möglichkeit die Tc-Aufnahme für gegatete Aufnahmen zu verwenden. Die Phantommessungen zeigen das Potential des Reverse-SDI Protokolls für die klinische Praxis. In folgenden klinischen Patientenstudien muss die Anwendbarkeit des Protokolls gezeigt und untersucht werden, ob die applizierte Tl-Dosis weiter reduziert werden kann um die Äquivalenzdosis von momentan 17,5 mSv auf 12,5 mSv (zum Vergleich 10–16 mSv nach EANM Empfehlung für Tc Protokolle) zu senken. Die sehr gute Übereinstimmung der Verhältnisse von Tc:Tl in den Phantommessungen, zeigen, dass die verwendete Methode auch zur Bestimmung im Patienten geeignet ist.

Literaturverzeichnis

1. Germano G, Berman DS. Clinical Gated Cardiac SPECT. Futura Inc.; 1999.
2. Botterweck H, Bippus R, Goedicke A, et al. Quantitative simultaneous multiple isotope SPECT imaging with iterative Monte-Carlo reconstruction. In: Proc. Fully 3D Image Reconstruction Meeting, Lindau; 2007. p. 221–4.
3. Backus BE, Verburg FA, Konijnenberg MW, et al. Myocardial perfusion SPECT: rest and stress in one acquisition. J Nucl Cardiol. 2008;15(4):S19.

3D Registration based on Normalized Mutual Information Performance of CPU vs. GPU Implementation

Florian Jung, Stefan Wesarg

Interactive Graphics Systems Group (GRIS), TU Darmstadt, Germany
`stefan.wesarg@gris.tu-darmstadt.de`

Abstract. Medical image registration is time-consuming but can be sped up employing parallel processing on the GPU. Normalized mutual information (NMI) is a well performing similarity measure for performing multi-modal registration. We present CUDA based solutions for computing NMI on the GPU and compare the results obtained by rigidly registering multi-modal data sets with a CPU based implementation. Our tests with RIRE data sets show a speed-up of factor 5 to 7 for our best GPU implementation.

1 Introduction

The registration of medical image data plays an increasing role in diagnosis as well as image-guided interventions. A widely used similarity measure for multi-modal registration is the normalized mutual information (NMI) metric [1]. In this work, we investigate to which extent the execution of an NMI based rigid registration can be sped up if parts of the algorithm are executed in parallel on graphics cards. A CUDA-based implementation on the GPU is compared to a CPU based solution for 3D rigid registration of medical image data.

By studying recent publications, attempts for employing GPU hardware and partially CUDA for solving image registration problems can be found. Köhn et al. [2] investigated the performance gain for rigid registration using a sum of squared differences metric, that performs well for intra-modality registration but has limitations for inter-modality registration [3]. Programming was entirely done on the GPU employing GLSL. An approach for 2D/3D registration based on automatic differentiation has been introduced by Grabner et al. [4]. It uses Cg for the GPU part of the algorithm, and a performance gain of factor 6 is reported. Muyan-Özcelik et al. [5] describe a CUDA implementation of Thirion's deformable registration algorithm [6] with a speed-up of 55 times compared to a CPU implementation. The problem of MI based registration using CUDA has been investigated by Shams et al. [7]. Their solution consists in an approximation of the probability density function, which is part of the MI based similarity measure calculation, employing a down-sampled joint histogram. In contrast to their work, we compute NMI based on histograms with a much higher number of bins for assuring a registration accuracy for the GPU approach that is equal to a CPU implementation.

2 Materials and Methods

The registration of two data sets requires their initial resampling in order to have isotropic voxels of equal resolution. This has to be done only once thus representing a negligible portion of the overall computation time. The other steps involved in our implementation are the interpolation of voxel data, the computation of NMI as similarity measure, an optimization step, and the application of a transformation to one of the data sets. They are executed iteratively until the optimization does not further improve the value of the similarity measure. The optimization step is done using a downhill simplex algorithm taken from the *Numerical Recipes* [8]. An initial test of the CPU reference implementation focused on the execution time of the remaining three steps. Obviously, the greatest speed-up can be achieved if steps which are computationally demanding are shifted to the GPU. Figure 1 shows the result of such an examination for the registration of a CT and an MRI data set of the head. It can be seen that NMI computation consumes most of the time. Thus, we decided to develop a CUDA implementation for this step.

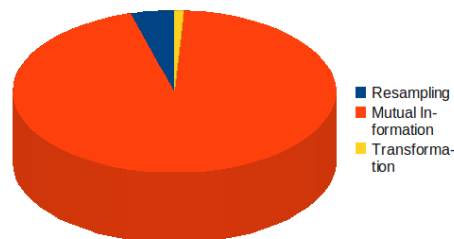
Employing an NMI metric requires the computation of two 1D histograms and one 2D joint histogram. For that, each voxel can be processed independently from its neighbors. In order to avoid that different threads concurrently increase the value of the same histogram bin, this requires mutual exclusion, a technique that is realized in CUDA with the `AtomicAdd()` function.

A straight-forward solution would be a direct port of the CPU implementation to CUDA. There, the number of threads would be equal to the number of voxels. It is obvious, that this will not result in fast code, since in case of 1D histograms with 256 bins which are stored in global memory only a maximum of 256 threads can write simultaneously, and this only if each of them accesses a different bin – which is rather unlikely in reality. In the following, we present two solutions with an increasing efficiency for the computation of 1D and 2D histograms as well as the calculation of the NMI metric. In our case, the 1D histograms have 256 bins and the 2D histogram is made up of 256×256 bins.

2.1 Independent Computation of 1D and 2D Histograms

We initialize each block with 256 threads – each of them computing the corresponding bins for one voxel position in both data sets. Each block holds

Fig. 1. The relative execution times for the steps being candidates for a CUDA implementation. It can be seen that NMI computation takes most of the time during a single iteration.



two 1D histogram copies in its corresponding shared memory, thus we benefit from the fast access to this type of memory. After initialization and filling of the 1D histograms in each block, each thread writes back the value of one bin from shared memory to the two resulting 1D histograms in global memory using `AtomicAdd()`.

Due to its limited size, it is not possible to store also the 2D histogram in shared memory. Therefore, we hold 256 additional copies of the 2D histogram in global memory, and each thread per block writes into a different 2D histogram. This approach requires an additional but very small ‘wrap-up’ kernel which consists of 256 blocks each with 256 kernels. Thus, each thread accumulates all values of the same 2D bin over the 256 2D histograms and writes the sum to the final 2D histogram – without any read/write conflicts.

2.2 Computation of 1D Histograms from the 2D Histogram

In fact, there is no need for the computation of the two 1D histograms if the 2D histogram is available: a simple projection of the 2D bin values to the two axes provides the information for the 1D histograms. In this second implementation, we compute in a first kernel the 2D histogram as described above. Afterwards, a second kernel is launched, where each thread computes the sum over all bins of one row and one column of the 2D histogram, respectively. For an optimal load balancing this second kernel contains 16 blocks each with 16 threads.

2.3 Interpolation and NMI Computation

The CPU reference implementation uses a nearest neighbor interpolation scheme due to its faster execution compared to a trilinear interpolation. In the case of GPU usage we could get a trilinear interpolation virtually for free. If texture memory is used instead of global memory for storing the image data on the graphics card, a fetch to the 3D texture memory delivers the trilinearly interpolated value. However, for the sake of comparability between CPU and GPU implementation, we use nearest neighbor interpolation also in the CUDA code.

The computation of NMI can be done efficiently using a reduction scheme. We employ code from the CUDA SDK (<http://www.nvidia.com/object/cuda>) vector reduction example that has been adapted to compute first the probability density functions (see [1]) followed by the reduction itself.

3 Results

The tests of our implementation have been two-fold. Firstly, we used two 8 bit head data sets from our own data base – one CT and one MRI data set – of the same patient for testing the behavior over different hardware generations. The CT data set contained $512 \times 512 \times 49$ voxels, the MRI data set $256 \times 256 \times 50$ voxels. On the CPU (Fig. 2) an increasing speed which each new processor version could

be noticed. It mainly scaled with the CPU's frequency, since we used a single-threaded implementation. Surprisingly, we could realize a lower performance of the CUDA implementation on graphics cards which are considered to be the fastest compared to the ones just below in the NVIDIA lineup (Fig. 2).

Secondly, we tested and compared the execution times on current hardware. CPU performance was measured on an Intel Quad Core Q6600 processor with 2.4 GHz (single-threaded code). As GPU an NVIDIA GeForce GTX 260 was chosen. For these tests, we used data sets from the open RIRE data base (<http://www.insight-journal.org/rire>). The results are shown in tab. 1. There, only the times for the most efficient (second) CUDA implementation are given revealing a speed-up of factor 5 to 7.

4 Discussion

We have presented new approaches for performing an NMI based 3D registration employing the GPU. With efficient CUDA implementations, a significant speed-up compared to a CPU implementation could be achieved. But, we compared our approach only to a single-threaded CPU solution. Employing multi-threaded programming on multi-core CPUs would reduce the performance gain of the GPU solution. Current limitation for a better performance on the GPU is the available hardware. Our best approach holds 256 copies of the 2D histogram in global memory due to the limited size of shared memory. If the latter one increases – as expected for one of the next generations of NVIDIA GPUs – each block could have its one 2D histogram for much faster read/write access.

Compared to an existing MI based registration approach [7], we use the full histograms – instead of a down-sampled version – for avoiding potential problems with registration accuracy. Here, our implementation is performing

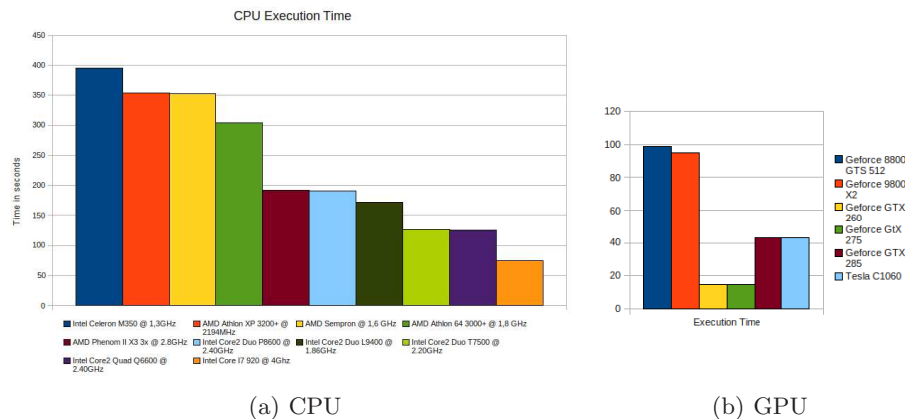


Fig. 2. The execution time for the registration of the CT and MRI head data set for different CPUs (a) as well as different GPUs (b). Note the different scaling of both diagrams!

Table 1. The execution times needed for the registration of image data sets from the RIRE data base (<http://www.insight-journal.org/rire/>).

RIRE patient	Time CPU	Time GPU	Speed-up
001: CT – MR T1	46.64 s	8.99 s	5.2
001: CT – PD	86.80 s	14.34 s	6.0
001: MR T1 – PET	64.18 s	9.45 s	6.8
001: CT – PET	42.17 s	7.43 s	5.7
003: CT – MR T1	65.14 s	9.37 s	7.0
006: CT – PET	79.26 s	11.06 s	7.2
007: CT – MR T1	68.38 s	10.41 s	6.6
008: PET – PD	47.23 s	7.46 s	6.3

better – Shams et al. note that using an exact histogram makes their GPU solution slower than the CPU – which is, to be fair, also due to the fact that we can use the AtomicAdd() function, which had to be simulated in software in the aforementioned work.

Finally, we could perceive the strange behavior of the GTX 285 (240 cores, 1.4 GHz) compared to the GTX 260 (216 cores, 1.2 GHz) – being more than 3 times slower. Discussions with other people in the CUDA community indicated that the reason for that are different memory partitions for both GPUs. Future work will focus on taking this into account for getting the best out of the latest GPU generations.

References

1. Pluim JPW, et al. Mutual-information-based registration of medical images: a survey. *IEEE Trans Med Imaging*. 2003;22(8):986–1004.
2. Köhn A, et al. GPU accelerated image registration in two and three dimensions. In: *Proc BVM*; 2006. p. 261–5.
3. Hill DLG, et al. Medical image registration. *Phys Med Biol*. 2001;46(3):R1–R45.
4. Grabner M, et al. Automatic differentiation for GPU-accelerated 2D/3D registration. *Lect Notes Computer Sci Eng*. 2008;64:259–69.
5. Muyan-Özcelik P, et al. Fast deformable registration on the GPU: a CUDA implementation of demons. In: *Proc ICCSA*; 2008. p. 223–33.
6. Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. *Med Image Anal*. 1998;2(3):243–60.
7. Shams R, et al. Speeding up mutual information computation using NVIDIA CUDA hardware. In: *Proc DICTA*; 2007. p. 555–60.
8. Press WH, et al. *Numerical recipes in C: the art of scientific computing*. 2nd ed. Cambridge University Press; 1992.

Experimental Assessment of Infarct Lesion Growth in Mice using Time-Resolved T2* MR Image Sequences

Nils Daniel Forkert¹, Dennis Säring¹, Andrea Eisenbeis², Frank Leypoldt³,
Jens Fiehler², Heinz Handels¹

¹Department of Medical Informatics,

²Department of Diagnostic and Interventional Neuroradiology,

³Department of Neurology,

University Medical Center Hamburg-Eppendorf, 20246 Hamburg

`n.forkert@uke.uni-hamburg.de`

Abstract. Cerebral ischemic stroke is a major reason for death in Germany and worldwide. Although lots of research has been done on long-time (3 hrs – 27 days) stroke growth, not much is known about the acute stroke growth. In this paper a method for the image based analysis and visualization of short-time infarct growth in murine model based on time resolved T2* MR image sequences is presented. After manual definition of initial healthy and infarct volumes the corresponding histograms are extracted and used for classification of the stroke tissues for every T2* image sequence. After automatic classification computed thresholds can be used to 3D visualize stroke growth dynamically and analyze the infarct size over time. The method proposed was validated by medical experts in terms of visual comparison to histological images of the brains after stroke. The medical experts reported that the accuracy of the projected infarct size is sufficient for further analysis steps of infarct growth. In summary the method proposed can help to understand the evolution of strokes and may help to get new insights in this disease and improve the therapy for stroke patients in future.

1 Introduction

Cerebral ischemic stroke is a major reason for death in Germany and worldwide. The ischemic stroke is caused by undersupply of blood and oxygen to the brain resulting from blockage of an artery due to blood clots or pieces of fatty deposits. It is characterized by insufficient metabolic circulation as well as reduction of cerebral blood flow and high oxygen extraction fraction. A detailed understanding of cerebral stroke evolution is clinically important for the diagnosis and treatment of patients [1]. The number of publications dealing with stroke growth in human as well as in animal models is high. In clinical practice several MR modalities, which reflect histopathologic changes in the brain tissue, are used including diffusion weighted imaging (DWI), apparent diffusion coefficient

(ADC) and T2-weighted measurements. Due to transportation time to hospitals, image sequences from humans directly after an occurrence of an occlusion are not available. Due to this fact most research focuses on the question how the stroke area grows over long time. Lansberg et al. [2] for example investigated time periods ranging from 19 hours to 27 days. The focus of this study is to investigate whether signal changes of T2* weighted MR images exhibit useful variations in the ischemic hemisphere immediately after vascular occlusion, such that the acute stroke growth can be analyzed. In this study a murine model has been utilized for short time infarct growth analysis. For the quantitative analysis of stroke growth a segmentation of the corresponding volume is required, whereas manual definitions of the infarct region have been incorporated in most studies (e.g. [3]). This manual definition is very time consuming and especially because of the high number of T2* image sequences acquired in this study such a procedure was not applicable. For this reason a semi-automatic method for quantitative and qualitative acute stroke evolution analysis based on T2* weighted MR image sequences was developed and is presented in the following.

2 Material and Methods

2.1 Stroke Model

Cerebral ischemia stroke was induced in male adult mice (n=18, 10-18 weeks) following [4]. Briefly, the animals were anesthetized intraperitoneally by ketamine (115,3 mg/kg, concentration of 25 mg/ml) and xylazine (4 mg/kg, concentration of 20 mg/ml) for all procedures. Temporary middle cerebral artery occlusion (tMCAO) was induced by the intraluminal filament model by inserting a nylon filament (6-0) with a 24 μm rounded tip via the external carotide artery into the internal carotide artery. The common carotide artery was ligated before. The filament was gently forwarded 9-10 mm measured from the bifurcation. The animal was immediately placed in a u-shaped animal rail and MR imaging was performed. If needed, animals were re-anesthetized intraperitoneally with half the above dose. After 30 or 60 minutes, for reperfusion, animals were removed from the scanner and the filament retracted 7-9 mm under visual inspection. Mice were replaced into the scanner and measurements continued. All animal experiments described were approved by the local ethics committee.

2.2 MR Image Sequences

Misery perfusion and acute ischemia affect magnetic resonance signal intensities due to changes of the blood oxygenation level [5]. To measure a net increase in local tissue deoxyhemoglobin the blood oxygenation level-dependent (BOLD) technique can be used. Especially T2* weighted imaging as an indicator of changes in the local concentration of deoxyhemoglobin based on the BOLD effect became important for the analysis of an acute cerebral ischemic infarction in the past. For this study among others dynamic T2* weighted imaging was performed

using a 3.0T Intera Philips MR device. The resulting dynamic T2* weighted image sequences were acquired using a repetition time of 1 second, an echo time of 9.21 ms, a flip angle of 25°, image in-plane image resolution of 0.23 mm × 0.23 mm and a slice thickness 0.3 mm. Using this setup 25 dynamic 3D images were acquired per mouse with a temporal resolution of approx. 2 min.

2.3 Stroke Analysis and Visualization

The first step of the semi-automatic method for short-time stroke growth analysis is the preprocessing of the available data. Due to saturation effects a linear increase of the intensity values over time can be observed. For reduction of this, intensity equalization using a histogram matching technique is performed first. Then the first time point of the 4D T2* weighted image sequence is defined as the baseline in which the core of the stroke is already represented by reduced intensities. In order to allow a precise analysis of stroke growth the core infarct volume of interest (VOI) is defined manually by medical experts. In doing so the user can manually select points in the orthogonal views which are connected using Bezier curves. Besides the infarct volume a healthy VOI is also required for the following analysis and can be defined in the same manner Fig. 1(a). These VOIs are then used to compute the corresponding histograms for the defined infarct and healthy volumes in the baseline. In general it can be expected that the infarct volume is represented by lower intensities than a healthy volume in T2* image sequences. Based on this assumption the following classification of the temporal stroke growth can be performed. For classification the histograms of the extracted VOIs are normalized Fig. 1(b,c). The infarct histogram is subtracted from the healthy histogram and for noise reduction the resulting representation is B-Spline approximated Fig. 1(d). This representation assigns positive values to voxels which are more likely to be part of the infarct volume and negative values to voxels part of the healthy brain tissues. Based on this representation the positive and the negative part are both divided into 8 bins of equal areas under the graph such that each intensity value can be assigned to

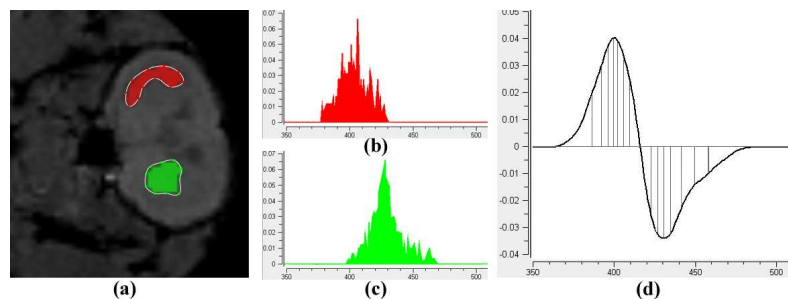


Fig. 1. Slice from a T2 MR image sequence with healthy (green) and infarct (red) VOI (a), corresponding histograms (b-c) and combined histogram representation divided into 16 bins (d).

one of the resulting 16 bins Fig. 1(d). This representation is then used for the analysis of stroke growth. For this the brain tissue is segmented in every available image. Normally the brain can be easily separated in T2* image sequences, such that a simple thresholding and a following largest connected component analysis leads to sufficient results. Based on the lower and upper thresholds of the computed bins, each voxel inside the brain segmentation is assigned to one of the computed bins Fig. 2. After this procedure is carried out for every time point the user can interactively chose which bins are supposed to be used for the dynamic 3D visualization of the stroke growth Fig. 3. The possibility of choosing different bins allows the visualization of different infarct areas such as the core or the penumbra. For visualization purposes the user defined infarct VOI is used to extract voxels part of the chosen bins for every time point. These voxels are then used for a seeded region growing in the bin-dataset, such that voxels outside the initial VOI can also be segmented. Finally this segmentation can be used for surface model generation using the Marching Cubes algorithm which can be displayed over time. Furthermore the extracted segmentation can be used for stroke quantification over time.

3 Results

The qualitative evaluation of the method proposed was validated by visual inspection by neuroradiologic experts. In a first step the semi-automatic computed final infarct volumes were compared to histological cuts of the corresponding brains of each dataset using different thresholds. Based on these thresholds, dynamic visualizations of all datasets were computed and inspected in regard to plausibility. In general the visualizations of the infarct growth were rated to display a feasible behavior. Especially the possibility using different bins for visualization was found to be helpful to investigate different stroke growth patterns. The visualizations were rated as helpful to get new insights into the short-time growth of ischemic strokes, especially the interactive navigation through time and 3D space, including rotation and zooming was judged as a benefit.

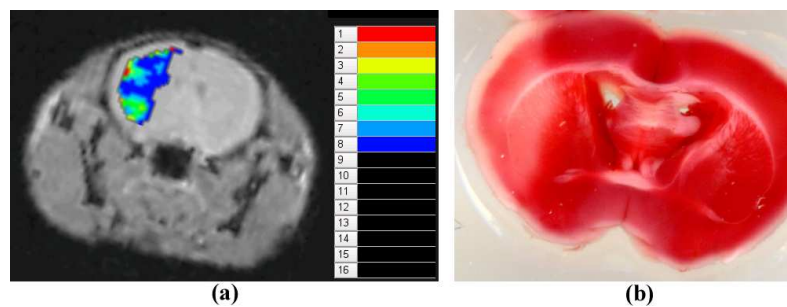
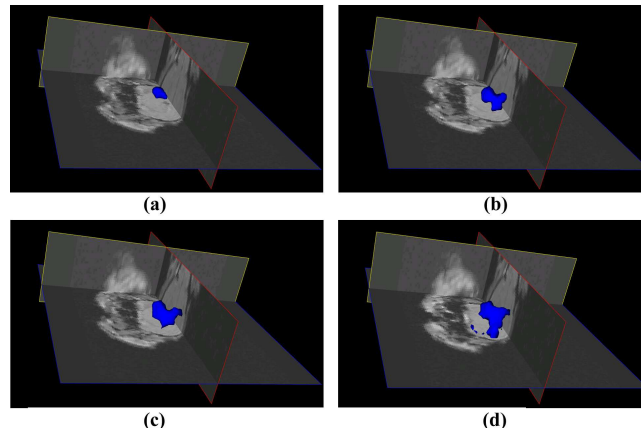


Fig. 2. Slice from a T2* MR image sequence and color-coded (bin 1-8) extracted infarct region (a) and corresponding histological slice from mouse brain (b).

Fig. 3. Selected frames of the dynamic visualization of acute stroke growth for the timepoints $t = 0\text{min}$ (a), $t = 18\text{min}$ (b), $t = 36\text{min}$ (c) and $t = 50\text{min}$ (d).



4 Discussion

In this paper a method for visualization and quantitative analysis of infarct growth was presented. Based on one initial healthy and infarct volume of interest the stroke can be segmented for every time point available using a histogram based classification and dynamically displayed over time. It is planned to include the information of the histological images for automatic VOI definition in order to reduce the manual interaction time. At the moment the method proposed is used for a clinical trial and first medical results will be published soon. Further more post-occlusion image sequences will be included in order to quantify the tissue which is reactivated after occlusion has been removed. In summary the method presented can help to explore alterations explicitly during the first few minutes of vascular occlusion and can help to understand the evolution of strokes.

References

1. Liu Y, D'Arceuil HE, Westmoreland S, et al. Serial diffusion tensor MRI after transient and permanent cerebral ischemia in nonhuman primates. *Stroke*. 2007;38(1):138–45.
2. Lansberg MG, Thijs VN, O'Brian MW, et al. Evolution of apparent diffusion coefficient, diffusion-weighted and T2-weighted signal intensity of acute stroke. *Am J Neuroradiol*. 2001;22:637–44.
3. Rosso C, Hevia-Montiel N, Deltour S, et al. Prediction of infarct growth based on apparent diffusion coefficients: Penumbra assessment without intravenous contrast material. *Radiology*. 2009;250:184–92.
4. Gelderblom M, Leypoldt F, Steinbach K, et al. Temporal and spatial dynamics of cerebral immune cell accumulation in stroke. *Stroke*. 2009;40(5):1849–57.
5. Roussel S, van Bruggen N, King M, et al. Identification of collaterally perfused areas following focal cerebral ischemia in the rat by comparison of gradient echo and diffusion-weighted MRI. *J Cereb Blood Flow Metab*. 1995;15:578–86.

Multimodale Registrierung von Knochen-Szintigraphien und Röntgenbildern

Nils Papenberg¹, Hanno Schumacher², Stefan Heldmann¹, Tobias Böhler³,
Dörte van Straaten³, Stefan Wirtz³

¹Fraunhofer MEVIS, Wallstraße 40, 23562 Lübeck

²MiE medical imaging electronics GmbH, Hauptstraße 112, 23845 Seth

³Fraunhofer MEVIS, Universitätsallee 29, 28359 Bremen

nils.papenberg@mevis.fraunhofer.de

Kurzfassung. Die Arbeit beschreibt eine neue Methode zur multimodalen Registrierung von funktionalen und morphologischen Bilddaten. Das Verfahren basiert auf einer speziellen Vorverarbeitung beider Bilder, die in beiden Daten enthaltene Informationen in Maskenbilder extrahiert. Diese werden mittels monomodalem Abstandsmaß in der Registrierung verglichen. Es wird dabei ein um Skalierung erweitertes rigides Transformationsmodell verwendet. Das Verfahren ist an planaren Bilddaten aus der Veterinärmedizin getestet und dabei als vielversprechend bewertet worden.

1 Einleitung

In dieser Arbeit präsentieren wir ein neues Registrierungsverfahren für nuklearmedizinische Daten und klassische Röntgenbilder. Funktionelle Bildgebung, wie die Szintigraphie, wird dazu verwendet, pathologische Prozesse wie Entzündungen oder Funktionsstörungen sichtbar zu machen, enthält aber wenig Informationen über die genaue Anatomie des dargestellten Gewebes. Daher werden solche Daten mit morphologischen Bildgebungen, die diese Anatomie deutlich sichtbar machen, kombiniert. Hier stellt sich das klassische Registrierungsproblem: Die Notwendigkeit der Überlagerung von funktionaler und anatomischer Bildgebung, um die funktionalen Informationen mit den anatomischen Positionen zu kombinieren. Diese Fragestellung tritt sowohl bei zweidimensionalen Bildern [1] als auch bei tomographischen Daten, wie PET-CT-Fusion [2], auf. Das Registrierungsproblem ist multimodal und wird zusätzlich durch das schlechte Signal-zu-Rausch-Verhältnis der nuklearmedizinischen Bilddaten erschwert. Wir bearbeiten in dieser Arbeit eine klinische Fragestellung aus der Veterinärmedizin: Bei lahmen Pferden kann zur Diagnose eine Szintigraphie der betroffenen Gelenke aufgenommen werden. Für den Veterinär von Interesse ist hierbei die Ansammlung der gemessenen Zerfälle (Counts) in der Nähe des pathologischen Gelenkspaltes, um Rückschlüsse über die Schwere der Erkrankung zu ziehen. Die Szintigraphie soll mit einer herkömmlichen Röntgenaufnahme des Gelenks kombiniert werden, um genaue anatomische Informationen zu erhalten.

Diese Arbeit befasst sich mit einem neuen Verfahren zur numerischen Lösung dieses multimodalen Registrierungsproblems. Wir zeigen, dass durch eine geeignete Vorverarbeitung der Datensätze das Problem auf ein monomodales Registrierungsproblem zurückgeführt werden kann. Das Verfahren ist nicht auf zweidimensionale Daten beschränkt und kann auch im Kontext humanmedizinischer Fragestellungen angewendet werden.

2 Registrierungsansatz

Die Bilddaten folgen bei der Untersuchung mehreren Regeln: Während beider Aufnahmen steht das Pferd jeweils vor der Bildgebung. In beiden Bildern ist das Gebiet aus der gleichen Richtung (vorne oder seitlich) abgebildet. In Abb. 1 ist ein Paar korrespondierender Daten dargestellt.

Der Ansatz der Registrierung basiert auf einer geeigneten Vorverarbeitung beider Modalitäten und der anschließenden Anwendung eines monomodalen Abstandsmaßes zur Bestimmung einer geeigneten Transformation, die abschließend wiederum auf die originären Daten angewendet wird. Ziel der Vorverarbeitung ist die jeweilige Generierung eines Maskenbildes, das die Lage der Knochen in den Daten beschreibt, da diese in beiden Modalitäten deutlich sichtbar sind. Die Vorverarbeitung erfolgt in mehreren Schritten:

1. Gaußglättung [3] zum Entrauschen der Daten,
2. Segmentierung der Knochen vom Hintergrund mittels Schwellwertverfahren [4],
3. morphologisches Closing [5] zum Schließen eventueller Lücken in der dargestellten Struktur.

Die Verarbeitung der Röntgendaten erfolgt analog, wobei auf ein vorgeschaltetes Glätten aufgrund der Qualität der Daten verzichtet wird. Die Ergebnisse der Vorverarbeitung sind in Abb. 1 dargestellt.

Aufgrund der Aufnahmeumstände der beiden Bilder lässt sich die zu erwartende Transformation zwischen den Daten auf Translation, Rotation und Skalierung beschränken. Hierbei kommt die Skalierung durch die unterschiedlichen

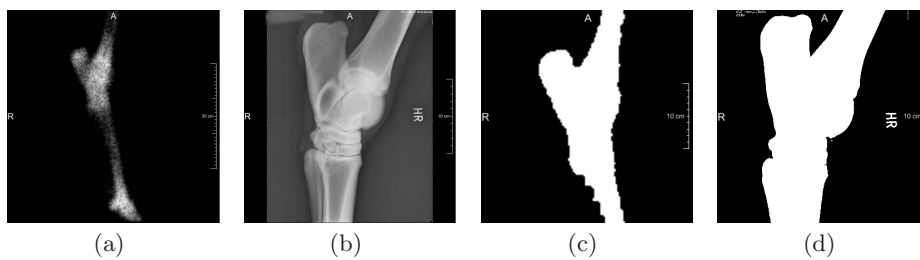


Abb. 1. Eingangsbilder Szintigraphie des gesamten Hinterbeines (a) und Röntgenbild des betroffenen Gelenks (b); Maskenbilder der Szintigraphie (c) und der Röntgenaufnahme (d).

Abstände des Tieres vom bildgebenden Gerät zustande. Daher verwenden wir in dieser Arbeit ein rigides Transformationsmodell. Dies ist um die Möglichkeit der Skalierung erweitert worden, da durch die unterschiedliche Entfernung des Tieres vom Aufnahmegerät die Größe der dargestellten Objekte in den Bildgebungen stark differieren. Für $(x_1, x_2)^T \in \mathbb{R}^2$ ist somit durch

$$\begin{pmatrix} y_1(x_1, x_2) \\ y_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \quad (1)$$

mit Skalierungsfaktor $s \in \mathbb{R}^+$, Rotationswinkel $\alpha \in \mathbb{R}$ und Translationsvektor $(t_1, t_2)^T \in \mathbb{R}^2$ der korrespondierende Punkt $(y_1, y_2)^T \in \mathbb{R}^2$ definiert. Da in der Szintigraphie eine größere Region abgebildet wird, wählen wir sie als Referenz und bestimmen durch das Verfahren die Lage des Röntgenbildes, das als Template genutzt wird, „in“ der Szintigraphie. Dies hat den weiteren Vorteil, dass wir nach der Registrierung mit der originären Szintigraphie weiterarbeiten können. Als Abstandsmaß für die Maskenbilder verwenden wir das SSD-Maß [6]. Mittels Gauß-Newton-Optimierung [7] wird eine rigide Transformation gesucht, die dieses Maß in Abhängigkeit der Parameter (α, s, t_1, t_2) minimiert. Zur Beschleunigung ist das Verfahren mit einer Multiresolutionsstrategie kombiniert worden [3].

3 Prototypische Umsetzung

Zur Anwendung und Validierung des oben beschriebenen Registrierungsverfahrens ist ein klinischer Prototyp entwickelt worden. Dieser bildet den gesamten Workflow des Veterinärs ab. Die Softwarelösung stellt zunächst Möglichkeiten für das Laden, Speichern und Visualisieren von korrespondierenden Datensätzen zur Verfügung. In einem zweiten Schritt wird die Registrierung voll automatisch ausgeführt. Anhand des Fusionsbildes und der darin enthaltenen anatomischen Information kann der Veterinär den zu untersuchenden Gelenkspalt manuell segmentieren. Daraufhin werden die Counts in der entsprechenden Region der Szintigraphie aufsummiert und dargestellt. Weiter stehen Standardwerkzeuge der Visualisierung (Grauwertfenstern, Zoomen, Farbänderungen, Verschieben) und die Möglichkeit des Abspeicherns der Ergebnisse zur Verfügung.

Der Prototyp wurde innerhalb der Rapid Prototyping Software MeVisLab [8, 9] erstellt. Die Registrierung wird mittels des MeVisLab-internen Registrierungsframeworks MERIT realisiert. Abbildung 2 zeigt einen Screenshot der Anwendung. In der oberen Hälfte der Anwendung werden die Daten fusioniert dargestellt, zusätzlich ist das Ergebnis einer Freihandsegmentierung gezeigt. Die untere Hälfte der Anwendung beinhaltet die Werkzeuge zur Bildmanipulation und zum Abspeichern der Ergebnisse.

4 Ergebnisse

Die Registrierung ist an sieben Datensätzen aus der Pferdeklinik Bargteheide getestet worden. Diese Datensätze zeigen unterschiedliche Ausschnitte der Tiera-

anatomie (Hals, Vorder- und Hinterbeine, Huf). Wie oben beschrieben stellt das Röntgenbild durch die genauere Fokussierung der Bildgebung jeweils einen kleineren Ausschnitt der Region dar. Die Datensätze zeigen Vorder- oder Seitenaufnahmen der Beine sowie Aufnahmen der Hufe von unten. Für alle Datensätze konnte durch die Registrierung eine Überlagerung der Daten erzielt werden. In Abb. 3, links ist dies exemplarisch für den in Abb. 1 gezeigten und einen weiteren Datensatz gezeigt.

5 Diskussion

Die gezeigten Ergebnisse sind von einem Veterinär als geeignet und vielversprechend beurteilt worden. Zur genauen Untersuchung der Eignung findet zur Zeit eine klinische Studie an der Pferdeklinik Bargteheide statt, die auch zur Evaluierung und Weiterentwicklung der prototypischen Software genutzt wird.

Die verwendeten Bildverarbeitungstechniken sind, jede für sich genommen, nicht neu, stellen aber in ihrer Kombination eine neue und effiziente Möglichkeit zur multimodalen Registrierung mittels eines monomodalen Maßes dar.

Das beschriebene Registrierungsverfahren ist nicht auf zwei-dimensionale Daten beschränkt und soll in einem nächsten Schritt als Grundlage zur Registrierung von SPECT- und CT-Volumendaten in der Humanmedizin dienen.

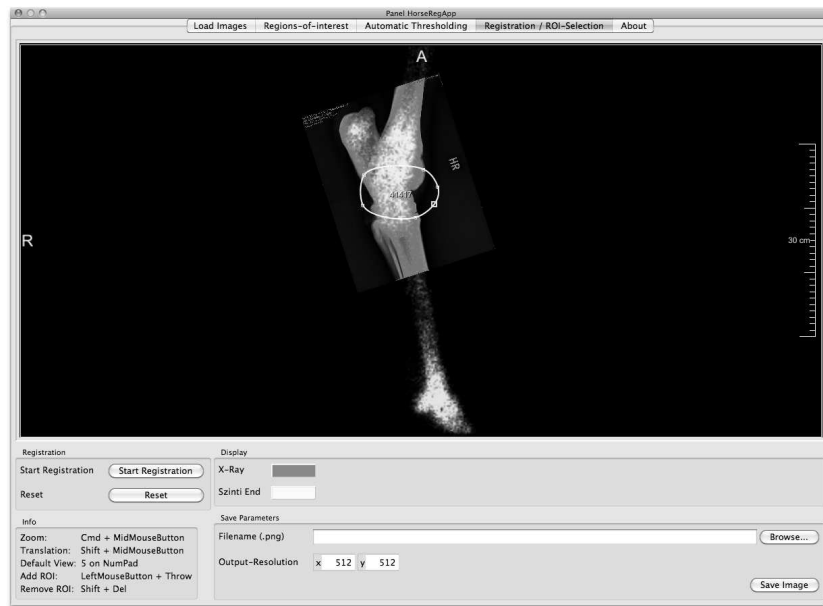
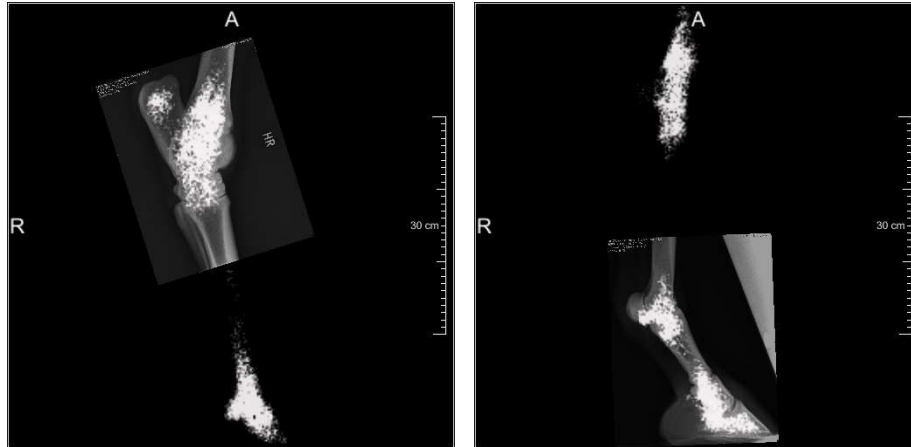


Abb. 2. Screenshot des Prototypen: Visualisierung der überlagerten Daten inklusive Segmentierung des Gelenks (oben); Werkzeuge zur Bildmanipulation und zum Speichern (unten).

Abb. 3. Ergebnisse der Registrierung: Überlagerung von Szintigraphie- (weiss) und Röntgenbild (grau) nach der Registrierung (links); Im zweiten Beispiel liegt der Focus der Untersuchung liegt auf dem Vorderhuf (rechts).



Danksagung. Die Autoren bedanken sich bei Dr. med. vet. W. Jahn und Dr. med. vet. V. Sill von der Pferdeklunik Bargtheide für die Bereitstellung und Begutachtung der Bilddaten.

Literaturverzeichnis

1. Currie GM, Wheat JM, Pearce RD. Planar image fusion. *Internet J Nucl Med.* 2007;4(1).
2. Mattes D, Haynor DR, Vesselle H, et al. PET-CT image registration in the chest using free-form deformations. *IEEE Trans Med Imaging.* 2003;22(1):120–8.
3. Jähne B. *Digitale Bildverarbeitung.* Springer, Berlin; 2005.
4. Handels H. *Medizinische Bildverarbeitung.* Vieweg und Teubner; 2009.
5. Soille P. *Morphological Image Analysis.* Springer, Berlin; 1998.
6. Modersitzki J. *Numerical Methods for Image Registration. Numerical Mathematics and Scientific Computation.* Oxford University Press; 2004.
7. Nocedal J, Wright SJ. *Numerical Optimization.* Springer, Berlin; 2006.
8. <http://www.mevislab.de>. MeVisLab.
9. Rexilius J, Kuhnigk JM, Hahn HK, et al. An application framework for rapid prototyping of clinically applicable software assistants. In: *Lect Notes Inform.* vol. 1; 2006.

Entwurf eines DICOM Structured Report am Beispiel Content-Based Image Retrieval

Petra Welter¹, Ralph Gülpers¹, Thomas M. Deserno¹, Jörg Riesmeier²
Marco Eichelberg³, Michael Onken³, Christoph Grouls⁴, Rolf W. Günther⁴

¹Institut für Medizinische Informatik, RWTH Aachen

²ICSMED AG, Oldenburg

³OFFIS - Institut für Informatik, Oldenburg

⁴Klinik für Radiologische Diagnostik, Universitätsklinikum Aachen

`pwelter@mi.rwth-aachen.de`

Kurzfassung. Der Einsatz von Digital Imaging and Communications in Medicine (DICOM) Structured Reporting (SR) ermöglicht einen standardisierten Austausch zwischen dem Picture Archiving and Communication System (PACS) und anderen klinischen Systemkomponenten. Ein strukturiertes Standardformat erlaubt die automatische Interpretation und eine Auswertung mittels Data Mining. Jedoch decken die im DICOM Standard festgelegten Structured Reporting Templates nicht alle Domänen ab. Wir zeigen beispielhaft das Design eines privaten SR-Templates für den Einsatz im Content-Based Image Retrieval (CBIR). Unser Lösungsansatz basiert auf dem Standard Computer-Aided Detection/Diagnosis (CAD) Template für Mammographie. Die Wiederverwendung bewährter SR-Module verspricht ein zuverlässiges Design. Wir analysieren die speziellen CBIR-Anforderungen und integrieren das neue Konzept von ähnlichen Bildern in unser Template. Allgemein anerkannte Templates für die Darstellung und den Austausch von Ergebnissen in einem standardisierten Format fördern den verbreiteten Einsatz neuer Verfahren in der klinischen Routine.

1 Einleitung

Der Nutzen neuer Verfahren in der medizinischen Informatik hängt wesentlich davon ab, wie gut sie in den Alltag eines Arztes integriert und angewendet werden können. Geht es um Ergebnisse, die ausgetauscht, präsentiert und dauerhaft gespeichert werden sollen, so ist die Verwendung eines Standardformats unabdingbar. Im klinischen Betrieb ist dies Digital Imaging and Communications in Medicine (DICOM). Für die Repräsentation von strukturierten Daten bietet DICOM das Structured Reporting-Format (SR).¹ Standard Templates² definieren zulässige Inhalte und Werttypen. Sie sind notwendig, um SR-Dokumente automatisch auswerten und systematisch nutzen zu können [1]. Neben z. Zt. ca. 35

¹ ftp://medical.nema.org/medical/dicom/2008/08_03pu.pdf

² ftp://medical.nema.org/medical/dicom/2008/08_16pu.pdf

allgemeinen Templates existieren spezielle für z.B. EKG und Röntgenstrahlendosis. Einige Templates können erweitert werden, falls dies für eine bestimmte Anwendung erforderlich ist. Ansonsten können neben den Standard-Templates auch private Templates definiert werden. Dies wird im Folgenden für die inhaltsbasierte Bildsuche (CBIR) durchgeführt. Der Nutzen von CBIR im medizinischen Bereich ist vielfach belegt [2]. Die Suche nach inhaltlich ähnlichen Bildern, ungeachtet ihrer textuellen Kategorisierung, und die damit ermittelten ähnlichen Patientenfälle, sollen als Diagnoseunterstützung eines Radiologen dienen.

2 Material und Methoden

2.1 Inhaltliche Anforderungen an das Template

Bei den inhaltlichen Anforderungen an ein CBIR-Template bedarf es u. a. der Angabe von Region of Interest (ROI). Um das Beispiel zu vereinfachen, beschränken wir uns auf folgende: a) ein Anfragebild, zu dem ähnliche Bilder zu suchen sind, b) Ergebnismenge der ermittelten ähnlichen Bilder, c) ihre jeweilige Ähnlichkeitsgüte und d) die Angabe des CBIR-Verfahrens.

2.2 Evaluierung existierender Templates

Im Bereich CAD finden sich unter den Standard Templates z. B. folgende:

- TID 2000 - Basic Diagnostic Imaging Report: Kann weder inkludiert noch erweitert werden. Es sieht selbst keine Beschreibung des CBIR/CAD-Verfahrens (d) vor. Die einzige Datenstruktur, die Anfragebild (a), Ergebnismenge (b) und Ähnlichkeitsgüte (c) aufnehmen könnte, ist *Diagnostic Imaging Report Elements*. Die Semantik für CBIR wird hier jedoch nicht deutlich.
- TID 4014 - CAD Image Quality: Hier ist nicht eine Ähnlichkeitsgüte (d) gemeint, sondern die Bildqualität, z. B. Stärke des Bildrauschens.
- TID 4000 - Mammography CAD Document Root: In dem inkludierten TID 4017 bzw. 4018 wird die Bildmenge genannt, auf der eine CAD Detektion bzw. Analyse durchgeführt wurde. Eine Ergebnismenge aller ähnlichen Bilder (b) kann nicht spezifiziert werden. TID 4017 und 4018 beinhalten Kennzeichnungen für bedeutsame Bildbereiche. Die Angabe einer ermittelten Ähnlichkeitsgüte (c) fehlt. Jedoch ist mit TID 4019 die Beschreibung zum CAD-Verfahren (d) aufgeführt.

2.3 Form und Aufbau einer Template Definition

Ein DICOM SR-Dokument enthält einen Dokumentenkopf (Header), wie ihn auch DICOM-Bilder besitzen. Er wird um einen Dokumentenbaum erweitert, der Struktur und Inhalt des eigentlichen Reports beinhaltet [3]. Templates beziehen sich auf den Dokumentenbaum. Sie werden als Tabelle spezifiziert. Jede Tabellenzeile definiert einen Baumknoten, der die Beziehung zum Vaterknoten (Rel)

Tabelle 1. Übersicht der Knotenbeziehungen in einer Template-Definition [1].

Nr.	Beziehung	Vaterknoten...
1	CONTAINS	enthält Komponente
2	HAS OBS CONTEXT	hat Beobachtungskontext
3	HAS CONCEPT MOD	hat Konzeptänderung
4	HAS PROPERTIES	hat Eigenschaft
5	HAS ACQ CONTEXT	hat Aufnahmekontext
6	INFERRED FROM	ist schlussgefolgert aus
7	SELECTED FROM	ist ausgewählt von

angibt, sowie Werttyp (VT), Konzeptnamen zur Erläuterung der Inhaltskomponente (Concept Name), Häufigkeit des Vorkommens (VM), Baumtiefe (NL) und ob der Knoten erforderlich ist (RT). In Tab. 1 sind mögliche Knotenbeziehungen erläutert. Ein SR-Dokument kann, beginnend ab dem Wurzelknoten des Dokumentenbaums, ein Root Template enthalten, welches wiederum durch Angabe von Template-Bezeichnern (TID) weitere Templates inkludiert.

2.4 Erzeugen eines SR-Dokuments

Bei der Erzeugung eines SR-Dokuments empfiehlt sich die Verwendung eines DICOM Toolkits mit entsprechender Unterstützung für die SR-Datenstrukturen. Wir verwenden DCMTK von OFFIS³. Die C/C++ Programmquellen lassen sich unter Windows, verschiedenen Unix-Betriebssystemen, sowie Mac OS X übersetzen. Ein alternatives bewährtes Open Source Toolkit [4] für Java ist z.B. dcm4che.⁴ DCMTK bietet mit dem Modul dcmsr Methoden für Zugriff, Lesen, Erstellen, Schreiben, Modifizieren, Drucken und Darstellen. Das folgende Beispiel für TID 4019 zeigt, wie ein neuer Knoten mit tieferer Verschachtelung hinzugefügt und der Name des Algorithmus angegeben wird:

```
srDoc->getTree().addContentItem(DSRTypes::RT_contains, \
    DSRTypes::VT_Text,DSRTypes::AM_belowCurrent);
srDoc->getTree().getCurrentContentItem().setConceptName \
    (DSRCodedEntryValue("111001", "DCM", "Algorithm Name"));
srDoc->getTree().getCurrentContentItem().setStringValue("IRMA");
```

3 Ergebnisse

Unser Template *CBIR_Root* (Tab. 2) beinhaltet in Zeile 2 das Anfragebild (a). Es inkludiert in Zeile 3 das Standard Template TID 4019 zur Angabe des CBIR-Systems (d) und in Zeile 3 das private Template *CBIR_Results* zur Angabe der

³ <http://dicom.offis.de>

⁴ <http://www.dcm4che.org>

Tabelle 2. SR-Template *CBIR_Root*.

Line	NL	Rel	VT	Concept Name	VM	RT
1			CONTAINER	EV(, , „CBIR Report“)	1	M
2	>	1	IMAGE	EV(, , „Query Image“)	1	M
3	>	1	INCLUDE	DTID(4019) „CAD Algorithm Identification“	1	M
4	>	1	INCLUDE	DTID(CBIR_Result) „CBIR Ergebnisse“	1	M

Tabelle 3. SR-Template *CBIR_Results*.

Line	NL	Rel	VT	Concept Name	VM	RT
1			CONTAINER	EV(, , „CBIR Results“)	1	M
2	>	1	CONTAINER	EV(, , „Scored Images“)	1-n	U
3	>>	1	IMAGE	EV(, , „Image“)	1	M
4	>>	1	NUM	EV(, , „Similarity Score“)	1	M

CBIR-Ergebnisse. *CBIR_Results* (Tab. 3) enthält in Zeile 2 einen Container für die Ergebnismenge, die sich in Zeile 3 aus einem Bild (b) und in Zeile 4 aus der zugehörigen Ähnlichkeitsgüte (c) zusammensetzt.

Die Erzeugung des vorgestellten Templates wurde als Applikation in C++ und DCMTK implementiert. IRMA [5], das CBIR-System unserer Wahl, übermittelt die CBIR-Ergebnisse an diese Applikation, welche das Template auf die Daten anwendet und ein DICOM SR-Dokument erstellt.

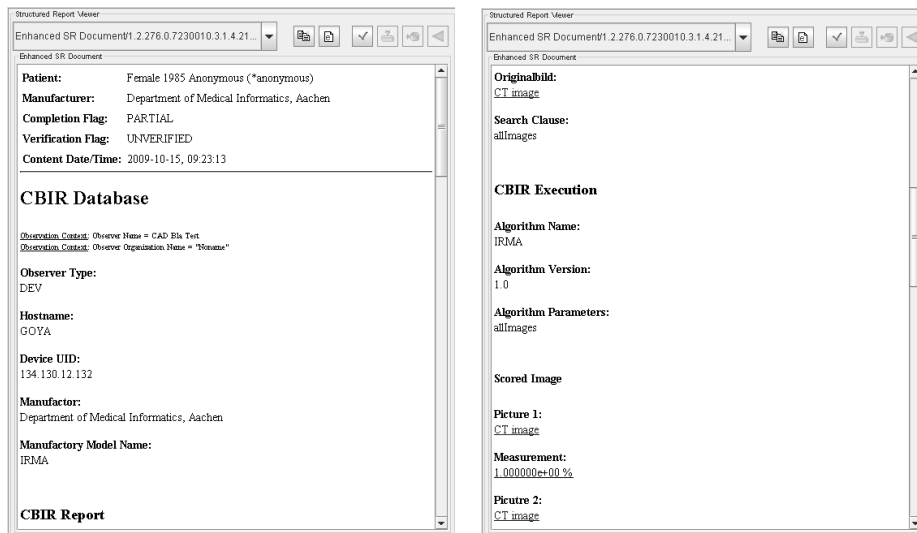


Abb. 1. Ausgabe des CBIR-SR in DICOMscope. Teile eins und zwei sind links und rechts dargestellt.

Zur Betrachtung des erzeugten SR-Dokuments verwenden wir DICOMscope von OFFIS. DICOMscope erzeugt wie die meisten SR-Viewer eine HTML-Ausgabe [6]. Das Layout ist allgemein gehalten (Abb. 1). Eine Konvertierung in das XML-Format mit dem Kommandozeilenprogramm *dsr2xml* erlaubt eine angepasste Darstellung über XSLT oder CSS. Alternativ ist mit dem Kommando *dsr2html* eine HTML-Konvertierung mit anschließender CSS-Anwendung möglich.

4 Diskussion

Dieser Beitrag bietet als Beispiel aus dem Bereich CBIR/CAD eine Hilfe bei der Erstellung privater SR-Dokumente. Die Verwendung strukturierter Reporte in einem standardisierten Format ermöglicht den zuverlässigen und einfachen Austausch von Ergebnissen. Dies fördert die Integration und damit den verbreiteten Einsatz von neuen Verfahren in der klinischen Routine.

SR-Dokumente werden u.a. in vordefinierten Arbeitsabläufen bzw. Profilen der Integrating the Healthcare Enterprise (IHE)⁵ eingesetzt. Für die Integration von CBIR in die radiologische Routine planen wir die Einbettung des vorgestellten SR-Templates im Post-Processing Profil, welches im Anschluss an die Bilderzeugung einsetzt.

Danksagung. Diese Arbeit wurde in Teilen von der Deutschen Forschungsgemeinschaft (DFG) gefördert, Le 1108/9

Literaturverzeichnis

1. Riesmeier J. Ein generisches Verfahren zur adaptiven Visualisierung von strukturierten medizinischen Befundberichten [PhD]. Universität Oldenburg. Oldenburg; 2006.
2. Müller H, Michoux N, Bandon D, et al. A review of content-based image retrieval systems in medical applications: clinical benefits and future directions. *Int J Med Inform.* 2004;73(1):1–23.
3. Eichelberg M, Aden T, Riesmeier J, et al. A survey and analysis of electronic healthcare record standards. *ACM Compute Surv.* 2005;37(4):227–315.
4. Vasquez A, Bohn S, Gessat M, et al. Evaluation of open source DICOM frameworks. *Proc MCIM Workshop*; 2007. p. C–10.
5. Lehmann T, Güld M, Thies C, et al. Content-based image retrieval in medical applications. *Methods Inf Med.* 2004;43(4):354–61.
6. Hussein R, Engelmann U, Schroeter A, et al. DICOM structured reporting part 2. problems and challenges in implementation for PACS workstations. *RadioGraphics.* 2004;24(3):897–909.

⁵ http://static.ihe.net/Technical_Framework/upload/ihe_tf_rev8.pdf

Validation of GEANT4 for Accurate Modeling of ^{111}In SPECT Acquisition

Bernd Schweizer, Andreas Goedicke

Philips Technology Research Laboratories, Aachen, Germany
bernd.schweizer@philips.com

Abstract. Quantitative SPECT imaging of ^{111}In is of growing importance in nuclear medicine applications like patient-specific radionuclide dosimetry. In this study, the quality of a Monte-Carlo based model of the imaging chain from nuclear decay to detection in a gamma camera is investigated. Important characteristics like the camera point spread function (PSF), the phantom attenuation and scatter effects are compared between simulation and experiment. The model shows a good agreement to the actual measurements and allows for further use in the development of quantitative SPECT reconstruction which will be mandatory for future oncology.

1 Introduction

Single Photon Emission Computed Tomography (SPECT) is an important nuclear imaging modality. In this study, imaging of ^{111}In is investigated, a radioisotope mainly used for applications in oncology. Here, the absolute assessment of local tracer activity provides access to both tumor viability staging and treatment prognosis. Due to its special characteristics including two equally dominant particle energy lines, ^{111}In imaging can become rather challenging in terms of image quality degradation e.g. due to down-scatter contamination. We present qualitative and quantitative comparisons between experimental and simulated gamma camera data acquisitions in order to prove the quality of a Monte-Carlo based imaging model to be used in the development of a next-generation quantitative SPECT reconstruction.

2 Materials and Methods

2.1 Phantom Acquisitions

Gamma camera acquisitions from two different phantom types have been acquired on a Philips Precedence SPECT/CT scanner at Rigs Hospital in Copenhagen. For PSF assessment, a planar measurement with a point source realized as an activity of 7.4 MBq ^{111}In located in a syringe tip was used. For the evaluation of attenuation and scatter impact, a SPECT scan from a NEMA PET

phantom was acquired (Fig.1(a)). Two cylinders have been filled with different concentrations of ^{111}In diluted in water: the outer cylinder (9.3 kBq/cm^3 , "warm background") and one insert (88.2 kBq/cm^3 , "hot insert"). In addition to that, two "cold" cylinders (air and Teflon) have been installed. Applied ^{111}In activities were determined with a Capintec CRC-15R dose calibrator. Volume measures were derived by a segmentation procedure based on CT images of the phantom. For the point source measurement, a source-to-camera distance of 15 cm was chosen and $5 \cdot 10^5$ counts were detected in a planar configuration. For the NEMA PET phantom, 128 projections with $(128\text{ pixels})^2$ resolution on a circular trajectory have been measured (Fig.1(a)). The acquisition time per projection was 30 s. In all acquisitions, a standard MEGP collimator was used and a standard dual ^{111}In energy window setting has been chosen for energy discrimination (1: 154 – 188 keV, 2: 221 – 270 keV).

2.2 Monte-Carlo Simulations

For Monte-Carlo simulations of the imaging chain, an advanced SPECT simulation framework based on GEANT4 [1] has been used, which was accelerated via the Forced-Detection-based variance reduction technique [2][3]. As an input for the simulation, voxelized activity and attenuation maps for the different phantom parts as well as the patient table models have been generated from the acquired CT-scan data (Fig.1(b)). The camera model included a standard MEGP collimator as well as a 9.5 mm scintillator crystal plate of NaI. Post-processing of the simulated detector events was performed to emulate binning impact, signal blurring (light spread) and effects of the detector electronics. According to the activity in the experiment, the number of simulated photon histories per projection was $2.4 \cdot 10^9$ for the point source and $4.4 \cdot 10^9$ for the NEMA PET phantom.

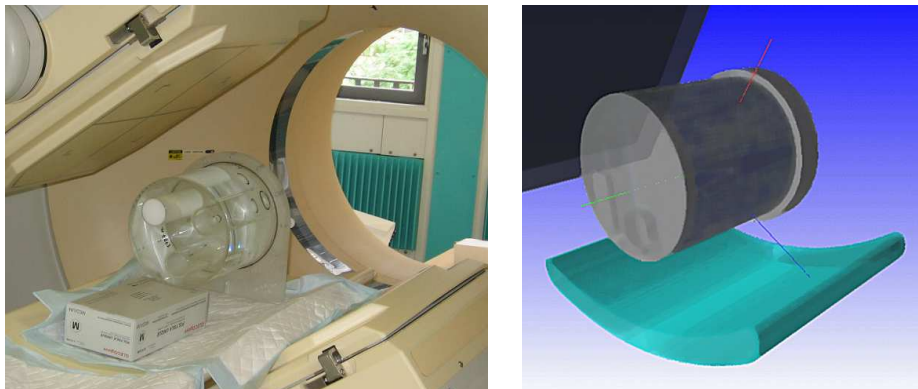


Fig. 1. Experimental setup for the NEMA PET phantom acquisition in the Philips Precedence SPECT/CT scanner (left). Modeling of the measurement in the GEANT4 based SPECT simulation framework (right).

2.3 Analysis

Planar gamma camera images of the point source were projected along pixel rows to obtain 2D profile plots of the PSF. The absolute camera sensitivity S was calculated from point-source images as $S = N/(A \cdot T)$, with N the number of detected counts, A the phantom activity and T the acquisition time. For the NEMA PET phantom, the detected counts in each projection were plotted against the projection number. The scatter fraction for each detection window was determined by applying a triple energy window (TEW) technique on the simulated projection data.

3 Results

The analysis of the PSF-profile plots (Fig.2) shows a good agreement between the experiment (red) and simulation of an idealized point source in air (green) with respect to signal intensity and shape in the central part of the curve. A further significant improvement, especially for the non-central areas was observed after including the patient tray and the camera lead shielding and plastic cover in the simulation (blue curve). The achieved improvement indicates that these signal components are due to additional scatter events from these parts of the simulation geometry. The remaining differences in the wing area of the curve can be explained by the basic strategy of the forced detection method to exclude those parts of the geometry from the radiation transport which generally only yield negligible contributions to the main result.

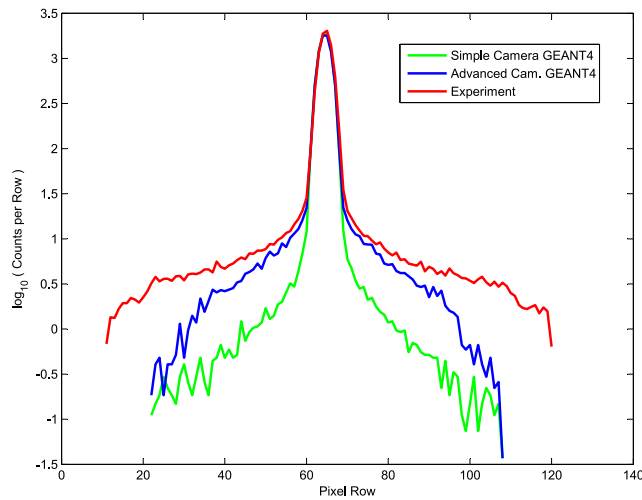


Fig. 2. Comparison between measured and simulated In-111 PSF on a logarithmic scale.

However, it should be noted that the signal level here is more than two orders of magnitude below the peak signal (and even lower, if the data is analyzed in the 2D detector area as opposed to a 1D-projection like in Fig.2). The agreement in the PSF amplitude is also reflected in the camera sensitivity figures, calculated

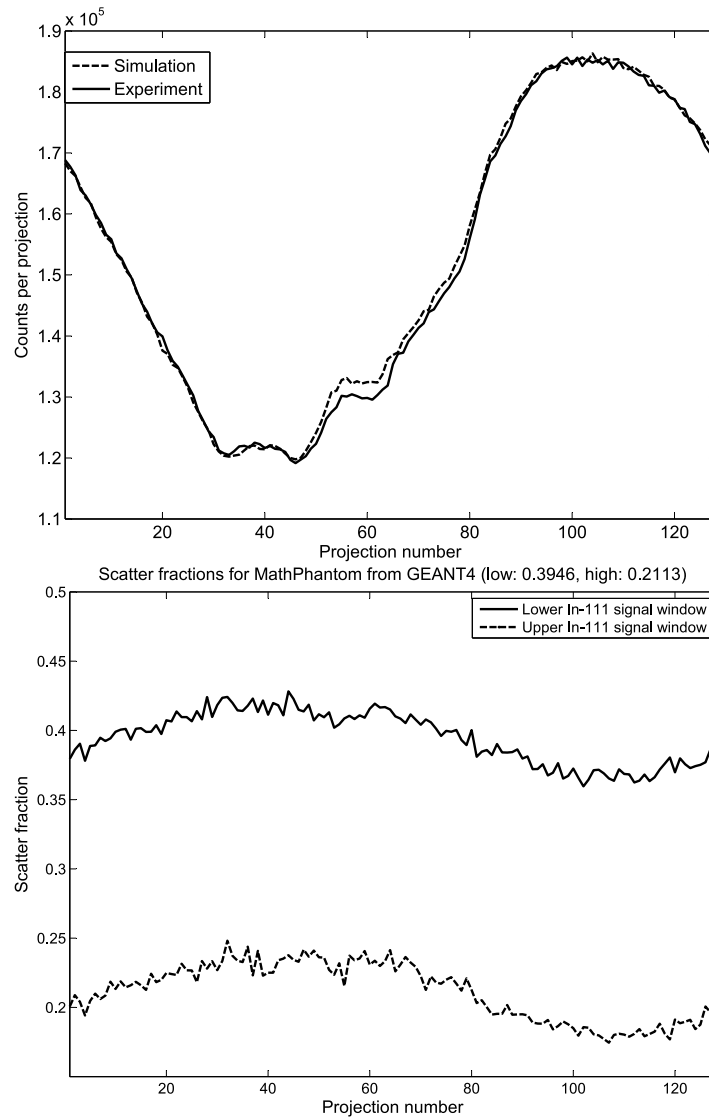


Fig. 3. Detected counts as a function of the camera angle for the NEMA PET phantom. Both curves are normalized to the same integral number of counts (top). Scatter fraction in lower and upper ^{111}In energy window as a function of the projection number for the NEMA PET phantom simulated in GEANT4 (bottom).

as $S = 163$ cps/MBq for the measured and 157 cps/MBq for the simulated point source acquisitions. In Fig.3(a), the detected counts per projection in the experiment and the simulation are compared with each other. Note that the focus of this analysis is on an analysis of the attenuation and scatter impact. Therefore the two curves were normalized to the same integral number of counts in order to facilitate the shape comparison. Obviously, the curves show an extremely good agreement. The determined integral difference between both plots before normalization was 6.9%, which is clearly within the boundaries of the experimental activity calibration error. In addition to that, Fig.3(b), displays the scatter fraction in the simulated signal for the NEMA PET phantom. The mean scatter fractions of 39.5% for the lower energy window and 21.1% for the upper energy window are in good agreement with other studies from literature [4], where for a similar approach values of 43% and 22% respectively are reported.

4 Discussion

The results achieved for both phantom types demonstrate a good agreement between experiment and simulation. In the PSF studies, the accurate modeling of important camera properties like PSF shape and absolute camera sensitivity could be demonstrated, especially when a realistic scanner model is included in the simulation geometry. This indicates that more complex phantom geometries can also be modeled accurately since such a case can be regarded as a superposition of point-sources from a mathematical view-point. In the NEMA PET phantom studies, detailed modeling of the complete imaging chain from nuclear decay to detection in the gamma camera could be proven. The quality of the Monte-Carlo simulation data warrants its use in quantitative SPECT. Future work based on these results will include the generation of highly realistic test datasets for algorithm development as well as use of accurate physical models in advanced SPECT reconstruction.

References

1. Agostinelli S, et al. GEANT4: a simulation toolkit. Nucl Instrum Methods Phys Res A. 2003;506:250–303.
2. de Jong HWAM, Beekman FJ, Slijpen TP. Acceleration of Monte Carlo SPECT simulation using convolution-based forced detection. IEEE Nucl Sci Symp Conf Rec. 1999 24-30 Oct;3:1532–6.
3. Goedicke A, Schweizer B, Staelens S, et al. Fast simulation of realistic SPECT projections using forced detection in GEANT4. In: Proc Med Biol Eng Conf; 2005.
4. Holstensson M. Optimisation of Window Settings for Quantitative ^{111}In Imaging: A Comparison of Measurements to Monte Carlo. Sweden: Lund University; 2006.

Features for Classification of Polyps in Colonoscopy

Sandy Engelhardt, Stefan Ameling, Stephan Wirth, Dietrich Paulus

Aktives Sehen, Computervisualistik, Universität Koblenz-Landau
engelhardt@uni-koblenz.de

Abstract. Colonoscopy is the gold standard for detection of colorectal polyps that can progress to cancer. In such an examination physicians search for polyps in endoscopic images. Thereby polyps can be removed. To support experts with a computer-aided diagnosis system, we compare different methods for automatic detection. Comparable to traditional pattern recognition systems, features are initially extracted and a classifier is trained on such data. Afterwards, unknown endoscopic images can be classified with the previously trained classifier. In this contribution we concentrate on the extension of the feature extraction module in the existing system. New detection methods are compared to existing techniques. Several features are tested, such as Graylevel Co-Occurrence Matrices (GLCM), Local Binary Patterns (LBP), and Discrete Wavelet Transform features. Different modifications on those features are applied and evaluated. We extend feature detectors to use color in different color spaces. We also compare different classifiers such as Support Vector Machines (SVM) and k -Nearest Neighbor classifier.

1 Introduction

Colonoscopy is the accepted gold standard for screening colon cancer or colorectal polyps, but there is a 6–12 % miss rate for adenomas that are 1cm or larger; the miss rate for smaller adenomas is up to 25 %. It is desirable to develop a system that marks polyps reliably during the screening process leading to a significantly decreased miss-rate.

Features for polyps are mostly from shape or from texture, or combined methods, as we outline in the following.

Krishnan et al. [1] approach is based on finding contours of abnormalities in the colon. The method of Hwang et al. [2] relies on the elliptical shape of colon polyps by applying watershed image segmentation. Wang et al. [3] propose a feature extraction method called Local Binary Pattern (LBP) which is a local texture descriptor. Karkanis et al. [4] propose a scheme which uses textural descriptors based on second order gray-level statistics called Graylevel Co-occurrence Matrices (GLCM), initially proposed by Haralick [5]. Karkanis et al. [6] propose a new color feature extraction scheme named Color Wavelet Covariance based on a fixed size sliding window. Further investigations are made here into these approaches.

All in all the results of these methods are hardly comparable because of the usage of different data bases, which are beside this often too small to make reliable predictions.

2 Materials and Methods

The data base in this contribution consists of four hours of video data from different colonoscopies initially used in [7]. It has been evaluated by medical specialists to obtain ground-truth data. Four scenes are extracted with polyps under varying illumination, view angle and distance. Each of the scenes has approximately 400 single frames and a resolution of 800×800 pixel. From the four scenes a heterogenous set of 130 frames is randomly chosen for further testings.

To represent ground-truth data image masks are created as depicted in Fig. 1. The white region describes the exact location, size and shape of a polyp.

The image is divided into overlapping patches. For each patch, a feature vector is computed and classified as *polyp* or *non-polyp* by considering the number of white pixels in the mask images, which must be larger than a certain threshold (≥ 625 pixel) to count as *polyp*. We extend [7] as we now use all patches and introduce new features.

The classification of the features is evaluated by the area under the ROC curve (AUC) value. Therefore, a data mining software called WEKA¹ and the classification library LibSVM² provide adequate means.

¹ <http://sourceforge.net/projects/wekaclassalgos>

² <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

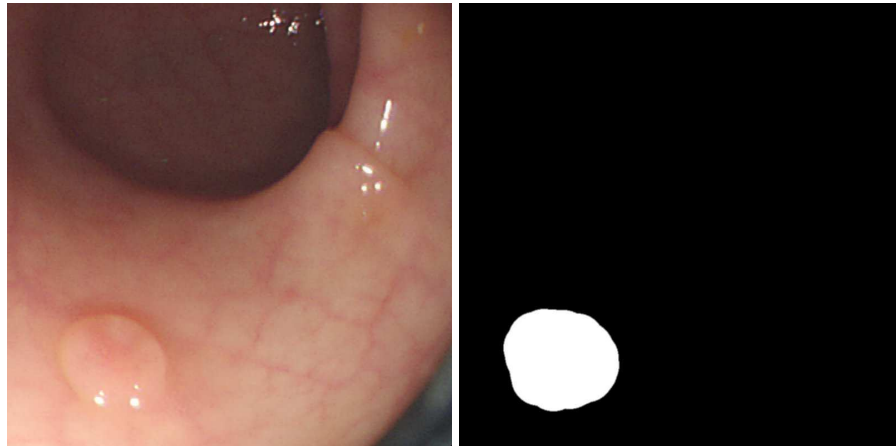


Fig. 1. Endoscopic image and its reference mask.

Table 1. Overview over the implemented features and their highest AUC value. Several properties are listed such as the dimensions of the feature vectors, color or gray information and patch size.

Feature	dim	color	gray	patch size	AUC LibSVM	AUC k -NN
ColorGLCM	72	×		32×32		0.843
L_8^{subset}	59	×		64×64	0.835	
ColorLBP	192	×		64×64	0.834	
OC-GLCM	108	×		64×64	0.832	
Color Wavelet	144	×		128×128	0.820	
OC-LBP	576	×		64×64	0.818	
WaveletDecomp	36	×		128×128	0.799	
$L_{16}^{\text{subset,ri}}$	54	×		64×64	0.799	
L_8^{ri}	108	×		64×64	0.792	
LBP	64		×	64×64	0.760	
GLCM6	6		×	64×64	0.740	
GLCM16	16		×	64×64	0.735	

We introduce new features based on texture such as ColorGLCM and ColorLBP, which are applications of GLCMs/LBPs on each RGB-color channel. The ColorGLCM method $F_m(P_{d,\theta}(I^i))$ produces for each channel i of the patch I four GLCMs $P_{d,\theta}$ with distance $d = 1$ and angle $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$. Then for each of the twelve GLCMs, six statistical measures m are extracted, namely energy, correlation, IDM, entropy, cluster shade and cluster prominence leading to a 72 dimensional feature vector. The Opponent-Color-GLCM (OC-GLCM) feature relates pairs of color channels by calculating GLCMs from the pixels of different color channels. Furthermore, several versions of rotation invariant LBPs and Subset LBPs [8] are tested.

Our wavelet features called ColorWavelet and WaveletDecomp also use GLCMs for representing texture after applying a three-level discrete wavelet transform to each patch and color channel. Again, statistical measurements are computed from the GLCMs which serves as input for the feature vector.

3 Results

Comparing the classification results from the SVM with the k -NN classifier leads to the conclusion that with less exceptions the SVM output has higher AUC values. No scheme could be investigated in which cases the latter performs better. Comparing the best classification results for each of the twelve features, the SVM holds the better results in eleven cases as shown in Tab. 1, which also clearly shows the advantages of using color. The best AUC result of all applied tests resulted from the ColorGLCM feature with a small patch size. The k -NN classifier performed best in this case with an AUC of 0.843.

4 Discussion

The published methods use very different image material. In this work, a very heterogenous set of images is chosen, containing frames from different scenes and different polyp types to make comparisons easier. We extend [7] by different texture descriptors and combine them to new features, incorporating wavelet transform, GLCMs and LBPs:

- Including color led to a significantly higher detection rate (+0.10 AUC for GLCM features). The single color methods performed equally well for the chosen data set. Only 0.05 AUC range lie between the best and the worst color method.
- The combination of all color channels of the RGB color space led to the best results.
- The discrete wavelet transform does not have the expected positive impact on polyp detection.
- The local binary pattern and the GLCM and their implemented variants perform equally well.
- The support vector machine classifier holds superior results in comparison to k -NN, considering the number of higher classification results.

Possible future extensions for a comprehensive polyp detection system are:

- Scale invariant features: Examine the resolution level of the texture by storing an additional parameter, which describes the distance to the intestinal wall during colonoscopy. An image normalization could be applied to achieve scale invariance.
- Over-complete wavelet transform: Compute translation invariant features by using the over-complete version of the wavelet transform (OCWT) [9].
- Preprocessing: Remove artifacts from endoscopic images such as shifted RGB color channels or glossy spots.
- Tracking: Prediction of a polyp's location to the next frame might help in classification.

References

1. Krishnan SM, Yang X, Chan KL, et al. Intestinal abnormality detection from endoscopic images. *Proc IEEE EMB*. 1998;2:895–8.
2. Hwang S, Oh J, Tavanapong W, et al. Polyp detection in colonoscopy video using elliptical shape feature. *Proc IEEE ICIP*. 2007;(2):465–8.
3. Wang P, Krishnan SM, Kugean C, et al. Classification of endoscopic images based on texture and neural network. *Proc IEEE EMB*. 2001;4:3691–5.
4. Karkanis SA, Magoulas GD, Grigoriadou M, et al. Detecting abnormalities in colonoscopic images by textural description and neural networks. *Proc Work Mach Learn Med Appl*. 1999; p. 59–62.
5. Haralick RM, Dinstein I, Shanmugam K. Textural features for image classification. *IEEE Trans Sys Man Cybern*. 1973 11;3:610–21.

6. Karkanis SA, Iakovidis DK, Maroulis DE, et al. Computer-aided tumor detection in endoscopic video using color wavelet features. *IEEE Trans Inf Technol Biomed.* 2003;7(3):141–52.
7. Ameling S, Wirth S, Paulus D, et al. Texture-based Polyp Detecion in Colonoscopy. In: *Proc BVM*; 2009. p. 346–50.
8. Mäenpää T, Ojala T, Pietikäinen M, et al. Robust texture classification by subsets of local binary patterns. In: *Proc. ICPR*; 2000.
9. Bradley AP. Shift-invariance in the discrete wavelet transform. In: *Proc Digit Image Comput Tech Appl*; 2003. p. 29–38.

Determination of a Vessel Tree Topology by Different Skeletonizing Algorithms

Andre Siegfried Prochiner¹, Heinrich Martin Overhoff²

¹Carinthia University of Applied Sciences, Klagenfurt, Austria

²University of Applied Sciences Gelsenkirchen, Gelsenkirchen, Germany
`heinrich-martin.overhoff@fh-gelsenkirchen.de`

Abstract. For navigated interventions of parenchymatous organs, plans must be adjusted according to organ motion and deformation. Vessel trees and especially their furcations are assumed to serve as landmarks for continuous registration. Three heuristic skeletonization algorithms were investigated to determine their ability to detect furcations correctly. A body-centered cubic grid-based algorithm was rated to fulfill these specifications, and an efficient implementation was developed.

1 Introduction

In navigated surgery and radiotherapy, interventions are often planned with respect to anatomical landmarks. Whereas distances between bony landmarks stay constant, landmarks in parenchymatous organs change their relative location e.g. due to altered patient positions or during resections. A vessel tree can serve as landmark system for such organs, because at least its topology remains – aside from resected parts – unaltered even during organ deformation, and vessel furcations can securely be detected.

Skeletonization is a common preprocessing step during the raster to vector transformation of a volume data set and consists of the reduction of objects in the 3D image volumes to a data reduced but topology conserving representation. For simplicity, a skeleton can be regarded as a medial axis of an object [1, 2]. More formally, it is the set of the centers of those spheres, which fully cover the object's volume and do not fully cover each other [3, 4].

Several algorithms, defined on a regular cartesian or on a body centered cubic grid respectively, were implemented in Matlab and investigated for their ability to generate correct skeletons of synthetic 3D objects. The most reliable one was chosen to skeletonize a vessel tree, that was constructed from a segmented 3D ultrasound image volume. Furthermore, the execution times of the programs were measured to detect hints for an efficient implementation.

2 Materials and Methods

In the following, a skeletonization is performed by removing voxels $\mathbf{p} \in \mathcal{O}$ from the object voxel set \mathcal{O} while preserving the object's topology.

A 3D object $\mathcal{O}^c \subset \mathbb{Z}^3$ with known skeleton $\mathcal{S}^c \subset \mathcal{O}^c$ was defined on a regular cartesian grid. This object was transformed into object \mathcal{O} and skeleton $\mathcal{S} \subset \mathcal{O}$ defined on a body-centered cubic (BCC) grid). Skeletonization algorithms were implemented for both grids:

- A the algorithm reported in [1] with modifications [5] for objects on a regular Cartesian grid,
- B the algorithm reported in [4] for objects on a BCC grid and
- C Algorithm B applied after distance transform.

Each algorithm generated a skeleton \mathcal{T}_i , $i = 1, 2, 3$. A visual evaluation of the coincidence generated vs. correct skeleton (\mathcal{T}_1 vs. \mathcal{S}^c , \mathcal{T}_2 and \mathcal{T}_3 vs. \mathcal{S}) was performed. In a second step, the algorithm with the best skeleton reconstruction was applied to generate the skeleton of the hepatic vein vessel tree.

2.1 Distance Transform

A distance transform determines the shortest distance d_i between a voxel $\mathbf{p}_i \in \mathcal{O}$ and the surface of this object. Voxels \mathbf{p}_i with identical distances $d_i = D_\kappa$ construct a voxel set \mathcal{S}_κ for $0 \leq D_\kappa \leq D_{\max}$, $0 \leq \kappa \leq \kappa_{\max}$. $\mathcal{O}_\kappa = \bigcup_{1 \leq j \leq \kappa} \mathcal{S}_j$ denotes the object's surface with thickness D_κ . Without the distance transform, all voxels \mathbf{p} of the surface of object \mathcal{O} are analyzed during skeletonization and can potentially be removed from the objects. When applying the transform, only the surface voxels \mathbf{p} having the current maximum surface distance D_κ are analyzed. The voxel decimation begins with $\tilde{\mathcal{O}}_0 = \mathcal{S}_0$. Due to the voxel decimation, surface voxels are eliminated from $\tilde{\mathcal{O}}$ which is thus transformed into $\hat{\mathcal{O}}$. Iteratively, for $1 \leq \kappa \leq \kappa_{\max}$ the decimation is performed on voxel sets $\tilde{\mathcal{O}}_\kappa = \hat{\mathcal{O}}_{\kappa-1} \cup \mathcal{S}_\kappa$.

2.2 Cartesian Grid Algorithm

In [5], an algorithm is presented where object surface points are classified by 62 neighborhood templates. If a voxel \mathbf{p}^c of the surface of object \mathcal{O}^c coincides with one of the templates, it can be removed.

2.3 Body-centered Cubic Grid Algorithms

A BCC grid is a subset of a regular cartesian grid and is defined by $\mathcal{B} = \{\mathbf{p} \in \mathbb{Z}^3 \mid p_x \equiv p_y \equiv p_z \pmod{2}\}$. Given an object $\mathcal{O} \subset \mathcal{B}$, the object's background is defined by $\tilde{\mathcal{O}} = \{\mathbf{p} \mid \mathbf{p} \notin \mathcal{O} \wedge \mathbf{p} \in \mathcal{B}\} = \{\mathbf{p} \mid \mathbf{p} \in \mathcal{B} \setminus \mathcal{O}\}$.

On a BCC grid, the distance vector sets \mathcal{D}_α are composed of the vectors \mathbf{v}_α to the nearest neighbors of a point $\mathbf{p} \in \mathcal{B}$ and the neighborhoods $\mathcal{N}_\alpha(\mathbf{p})$ are given by

$$\begin{aligned} \mathcal{D}_6 &= \{\mathbf{v}_6 \in \mathbb{Z}^3 \mid |v_x| + |v_y| + |v_z| = 2\} \\ \mathcal{D}_8 &= \{\mathbf{v}_8 \in \mathbb{Z}^3 \mid |v_x| + |v_y| + |v_z| = 3\} \\ \mathcal{D}_{14} &= \mathcal{D}_6 \cup \mathcal{D}_8 \end{aligned} \tag{1}$$

$$\mathcal{N}_\alpha(\mathbf{p}) = \{\mathbf{q} \mid \mathbf{q} \in \mathcal{B} \wedge \mathbf{v}_\alpha = \mathbf{p} - \mathbf{q} \in \mathcal{D}_\alpha\}, \quad \alpha = 6, 8, 14. \quad (2)$$

A point set $\mathcal{P} \subset \mathcal{Q} \subset \mathcal{B}$ is a connected component of \mathcal{Q} , if for each pair of points $\mathbf{p}, \mathbf{q} \in \mathcal{P}$ a sequence of points $\mathbf{p}_1, \dots, \mathbf{p}_m \in \mathcal{P}$ exists, such that $\mathbf{p}_1 = \mathbf{p}$, $\mathbf{p}_m = \mathbf{q}$ and $\mathbf{p}_i \in \mathcal{N}_{14}(\mathbf{p}_i)$ holds for $1 < i < m$. $C(\mathcal{Q})$ is the number of connected components of \mathcal{Q} . An object point $\mathbf{p} \in \mathcal{O}$ is a simple point, if

$$C(\mathcal{N}_{14}(\mathbf{p}) \setminus \mathcal{O}) = 1 \wedge C(\mathcal{N}_{14}(\mathbf{p}) \setminus \bar{\mathcal{O}}) = 1. \quad (3)$$

Simple points can be removed from object \mathcal{O} with conservation of the object's topology. In addition to pure skeletonization, in each step, a removed voxel can be assigned to its nearest remaining voxel. This assignment stops at skeleton voxels. Thus, the voxels each skeleton arm represents are known explicitly.

The classification of an object point \mathbf{p} as simple is the most frequently processed step and should be done efficiently. Two alternative simple point qualifications were implemented for Algorithm B as well as for algorithm 3:

- like in [4], the connectivity of 24 so-called checking rings $\mathcal{R}_{ij}(\mathbf{p}) \subset \mathcal{N}_{14}(\mathbf{p})$, $1 \leq i \leq 3, 1 \leq j \leq 8, |\mathcal{R}_{ij}| = 6$ was determined during skeletonization.
- The simplicity of a point \mathbf{q} was determined once for all of its $N = 2^{14}$ possible $\mathcal{N}_{14}(\mathbf{q})$ -neighborhoods and the binary results were stored in a database before skeletonization. During skeletonization, the simplicity of the individual $\mathcal{N}_{14}(\mathbf{p})$ -neighborhood was determined by data base access.

For large objects with irregular surfaces – like a highly resolved vessel tree volume – overskeletonization is a common problem. To avoid it, object surface's voxels \mathbf{p} are classified as stubble points, if $\mathbf{p} \in \hat{\mathcal{O}}_{\kappa-1} \wedge q = \mathcal{N}_{14}(\mathbf{p}) \cap \hat{\mathcal{O}}_\kappa \wedge |\mathcal{N}_{14}(\mathbf{p}) \cap \hat{\mathcal{O}}_\kappa| > 2$. Stubble points typically generate only short skeleton arms and are removed from $\hat{\mathcal{O}}_\kappa$ before the next simple point classification is processed.

3 Results

The skeletonization results are demonstrated in the following figures. Figure 1 shows a synthetic F-shaped object which consists of 313 voxels and its skeleton (opaque voxels) found by Algorithm C. Comparison with the ideal skeleton

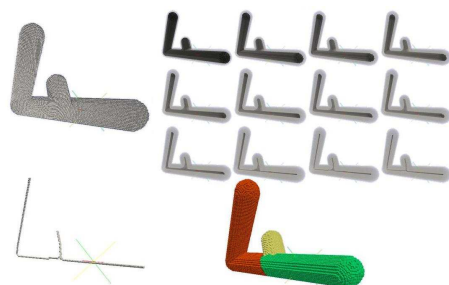


Fig. 1. Process of skeletonization and reconstruction of a synthetic object. (top left) Object, (top right) sequence of thinning steps with removed simple points (semi transparent) and remaining points (opaque), (bottom left) skeleton, (bottom right) voxels assigned to skeleton branches.

Table 1. Execution times for different realizations of simplicity testing (BCC algorithms).

simplicity test	single voxel	object	
		vessel tree 1	vessel tree 2
procedural	5.5 ms	1407 s	45178 s
data base	3.8 ms	1300 s	44874 s

(semi-transparent voxels) is performed by the euclidean distances between of the pairwise nearest voxels of ideal vs. real skeleton: 228 (80/0/3/2) voxels had distance 0 (1/1.41/1.73/2).

Algorithms A and B (Fig. 2) show results with additional furcations and reduced precision at the skeleton's branches. Obviously, the distance transform supports the detection of the correct skeleton.

Algorithm C was chosen to skeletonize two hepatic vein vessel trees. The trees were acquired in ultrasound image volumes and contained the inferior vena cava, the three branches of the hepatic vein and some of the segmental veins. Due to the limited image volume, the vessels are partially cut off. Anatomically, vessel trees 1 and 2 are identical, but tree 2 has the 8-fold resolution of tree 1. Visually, the skeletons reflect the topological properties well, especially for the segmental veins. In detail, for vessel tree 1, the left and right hepatic vein are marginally overskeletonized (Fig. 3). Each of the skeleton's arms belongs to one vein, i.e. none was erroneously assigned to two veins and no furcation was missing.

The Matlab code was processed on a PC equipped with an Intel Core2Duo E8400 3 GHz processor, 3 GB RAM, 1333 MHz main board FSB frequency. The execution times for the simplicity analysis are shown in Tab. 1 demonstrating that data base access is superior to procedural determination.

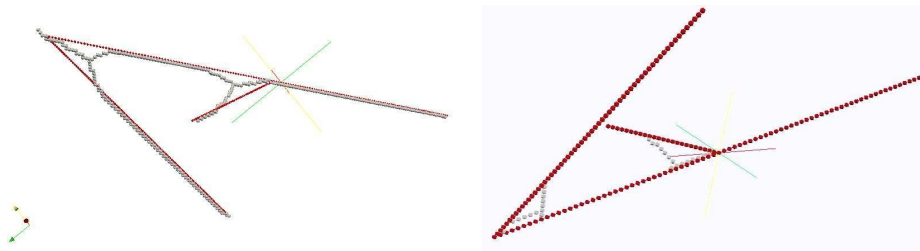
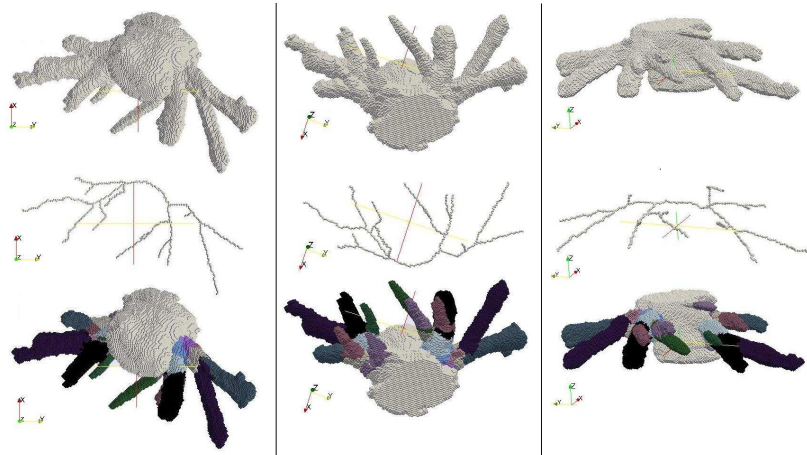


Fig. 2. Skeleton of the synthetic object (gray voxels) generated by (left) BCC grid based Algorithm B without distance transform and (right) cartesian grid based Algorithm A. Ideal skeleton is shown for comparison (red cubes).

Fig. 3. Hepatic vein vessel tree (top), its skeleton from Algorithm C (center), and its reconstruction (bottom). Distal (left column), proximal (center column) and right oblique (right column) view.



4 Discussion

The heuristics reported in [4] for objects defined on a BCC grid is rated as promising for the skeletonization of vessel trees. Simplicity analysis is faster using data base access than using explicit procedural analysis. Obviously, the skeletonization's execution times depend severely on the cardinality of \mathcal{O} .

Ongoing work is addressed to analyze the observed overskeletonization of vessels and to investigate the coincidence of vessel tree skeletons for data sets representing the same object, but but with reduced cardinality.

References

1. Ma M, Sonka M. A fully parallel 3d thinning algorithm and its applications. *Comput Vis Image Underst.* 1996;64:420–33.
2. Palagyi K, Kuba A. A 3D 6-subiteration thinning algorithm for extracting medial lines. *Pattern Recognit Lett.* 1998;19:613–27.
3. Blum H. A transformation for extracting new descriptors of shape. In: Wathen-Dunn W, editor. *Models for the Perception of Speech and Visual Form.* Cambridge: MIT Press; 1967. p. 362–80.
4. Brunner D, Brunnett G, Strand R. A high-performance parallel thinning approach using a non-cubic grid structure. *Chemnitzer Informatik-Berichte.* 2006; p. 1–13. Available from: <http://www.tu-chemnitz.de/informatik/service/if-berichte/pdf/CSR-06-08.pdf>.
5. Wang T, Basu A. An improved fully parallel 3D thinning algorithm. Univ. of Alberta, Dept. of Computer Science; 2005. <http://www.cs.ualberta.ca/TechReports/2005/TR05-31/TR05-31.pdf>, rev. 2009.09.05.

Online Detection of Straight Lines in 3-D Ultrasound Image Volumes for Image-Guided Needle Navigation

Heinrich Martin Overhoff, Stefan Bußmann

University of Applied Sciences Gelsenkirchen, Gelsenkirchen, Germany
heinrich-martin.overhoff@fh-gelsenkirchen.de

Abstract. 3D ultrasound imaging can be a valuable tool in navigated interventions. When used as continuous imaging device, it can not only show the anatomical site but has the potential to measure instrument positions, e.g. to validate the data of position localizer tools. Real time image analysis is a prerequisite for this concept. Needle-like instruments can be idealized by lines. An algorithm was implemented in C that determines four parameters of an equation of a line in 3D space from voxel positions. It mainly consists of a sequence of two 2-D Hough transforms, which determine the instrument's orientation and location. The execution time of the algorithm's C implementation was measured on a PC platform for varying numbers of instrument and error-prone voxels. For a moderate number of voxels, real time demands for the 4DoF position determination can be fulfilled.

1 Introduction

Planar (2-D) sonography is a well established online imaging modality to visualize anatomical structures during interventions like biopsies or punctures. Typically, a needle guiding device is fixed to the ultrasound transducer and aligned to the ultrasound image plane. The needle is oriented oblique to the image plane and therefore only partially visible in the sonograms. To navigate the needle during an intervention, image plane, anatomical structures and landmarks (e.g. vessels) must be properly aligned.

Interstitial high dose rate brachytherapy is a local radiooncologic therapy where radioactive sources are temporarily placed in catheters, that are themselves inserted using needle-like devices, so-called applicators [1]. Pre-interventional ultrasound imaging is increasingly used to detect and delineate organs and tumors for isodose planning in brachytherapy. Usually, the applicator placement is done "blind", without online imaging, but this can potentially be overcome using intra-interventional spatial (3D) ultrasound imaging. Reflections of anatomical structures like connective tissue or muscles can interfere with the applicator's reflections. Thus, adapting only the opacity transfer function when visualizing such an ultrasound image volume might neither be sufficient for the visualization nor for measuring the applicator's position.

For determining the location, orientation and tip (5 degrees of freedom (DoF)) of a needle-like instrument relative to an ultrasound image volume [2], an electromagnetic position measurement system is an option. In a clinical setup with 3D ultrasound imaging, an image based instrument pose determination can be of interest to validate location and orientation measurements (4 DoF). Such image volumes are analyzed by a new algorithm which was implemented in common C. The running time of that C program on standard hardware is measured to evaluate if real time demands are fulfilled.

2 Materials and Methods

Image volumes were acquired with a BC431E motorized 3D curved array transducer attached to a MyLab70 ultrasound device (Esaote Europe B.V., Maastricht, The Netherlands). An image volume consisted of approximately 250 images of 420×480 gray-scale pixels. For continuous acquisition, an acquisition rate of $r = 0.25$ volumes/second was assumed, resulting in a data rate of r volumes/second \cdot 250 images/volume \cdot $420 \cdot 480$ pixels/image \cdot 8 bit/pixel \approx 384.5 Mbit/second. Assuming additional processing times, e.g. for data transfer, image pre-processing, visualization and operating system activities, an execution time per volume $T_{\text{exec}} \leq 1$ second $\ll \frac{1}{r}$ was determined to fulfill real time demands.

For determining the applicator's pose, it is assumed that the applicator does not bend during the interactive feed motion. Thus, it is possible to presume that the reflections of the needle are represented by straight lines in all image volumes. The number of needle reflections or needle points \mathbf{r} respectively and the number of additional noise points varies in ultrasound image volumes, e.g. due to variable imaging properties or the needle length recorded in the image volume.

Such needle points \mathbf{r} fulfill the vector form of the equation of a line

$$\mathbf{r} - \mathbf{r}_0 - \lambda \mathbf{u} = \mathbf{0} \quad (1)$$

where \mathbf{u} denotes the direction and \mathbf{r}_0 the origin point vector. For an identical line, there is an infinite number of such representations. The equation of a line becomes unique, if the two conditions

$$\mathbf{u}^T \cdot \mathbf{u} = 1 \quad (2)$$

$$\mathbf{u}^T \cdot \mathbf{r}_0 = 0 \quad (3)$$

are fulfilled, i.e. if the length of the direction vector \mathbf{u} is normalized and origin point vector \mathbf{r}_0 and direction vector \mathbf{u} are perpendicular. Under these conditions, the equation of a line (Eq. 1) can be transformed into the pair of equations

$$\mathbf{u} \times (\mathbf{r} - \mathbf{r}_0 - \lambda \mathbf{u}) = \mathbf{u} \times (\mathbf{r} - \mathbf{r}_0) = \mathbf{0} \quad (4)$$

$$\mathbf{u}^T \cdot (\mathbf{r} - \mathbf{r}_0 - \lambda \mathbf{u}) = \mathbf{u}^T \cdot \mathbf{r} - \lambda = 0 \quad (5)$$

It contains five parameters: two components u_x and u_y of the direction vector \mathbf{u} , two components r_{0x} and r_{0y} of the origin point vector \mathbf{r}_0 which together define the line as a whole, and λ , which defines an individual point on the line.

A sequence of two 2-D Hough transforms was chosen to reduce the complexity compared to a 4-D Hough transform [3, 4]. The following algorithm was implemented to determine the four line parameters u_x , u_y and r_{0x} , r_{0y} for N_r scattered, noisy line voxels and additional N_{noise} noise voxels:

1. Direct binarization of the image volume to get N voxels $\mathbf{r}_i, 1 \leq i \leq N$, $N = N_r + N_{\text{noise}}$ being candidates for applicator reflections.
2. 2-D Hough transform for the \mathbf{u} -components: $\mathbf{u} \times (\mathbf{r}_i - \mathbf{r}_0) - \mathbf{u} \times (\mathbf{r}_j - \mathbf{r}_0) = \mathbf{u} \times (\mathbf{r}_i - \mathbf{r}_j) = -(\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{u} = \mathbf{0}$, $1 \leq i \leq N - 1, i < j \leq N$ and (Eq. 2) determine \mathbf{u}

$$-(\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{u} = \begin{bmatrix} 0 & r_{iz} - r_{jz} & r_{jy} - r_{iy} \\ r_{jz} - r_{iz} & 0 & r_{ix} - r_{jx} \\ r_{iy} - r_{jy} & r_{jx} - r_{ix} & 0 \end{bmatrix} \cdot \mathbf{u} = \mathbf{R}_{ij} \cdot \mathbf{u} = \mathbf{0} \quad (6)$$

Since $\text{rank}(\mathbf{R}_{ij}) = 2$, e.g. the third line of the system (Eq. 6) is substituted by (Eq. 2) to make it uniquely solvable. This leads to a non-linear equation system. If $u_z < 0$, $\mathbf{u} = -\mathbf{u}$. Component pair (u_x, u_y) is incremented.

3. Select the component pair (u_x^*, u_y^*) with maximal number of increments and calculate the optimal direction vector \mathbf{u}^* with (Eq. 2) for $u_z \geq 0$.
4. 2-D Hough transform for the \mathbf{r}_0 -components: With (Eq. 4) and $\mathbf{u} = \mathbf{u}^*$, the linear equation system for \mathbf{r}_0 follows

$$\mathbf{u} \times \mathbf{r}_0 = \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix} \cdot \mathbf{r}_0 = \mathbf{U} \cdot \mathbf{r}_0 = \mathbf{u} \times \mathbf{r}_i = \mathbf{b}_i = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad (7)$$

Since $\text{rank}(\mathbf{U}) = 2$, e.g. the third line of the system (Eq. 7) is substituted by (Eq. 3) to make it uniquely solvable, where the component pair (r_{0x}, r_{0y}) is incremented.

$$\begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ u_x & u_y & u_z \end{bmatrix} \cdot \mathbf{r}_0 = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} \quad (8)$$

5. Select the component pair (r_{0x}^*, r_{0y}^*) with maximal number of increments and calculate optimal direction vector \mathbf{r}_0^* with (Eq. 3).

Parallel needles have identical direction vectors \mathbf{u} . Parallel lines are discriminable, if \mathbf{r}_0 is noted in a rotated coordinate system where the z -axis is parallel to \mathbf{u} . In this case, different parallel needles have different (r_{0x}, r_{0y}) -pairs due to (Eq. 3), and the number of cluster points is equal to the number of needles.

3 Results

For a PC equipped with an Intel Core2Duo E8400 3 GHz processor and 3 GB RAM, the main board FSB frequency was 1333 MHz, and the C code was compiled on Windows XP Professional using Visual Studio 2008 with the SSE 2 instruction-set enabled. The processing times of the resulting program were measured for different numbers N of voxels \mathbf{r} for a resolution of $\Delta u_x = \Delta u_y = 0.01$. As table Tab. 1 shows, the running time is mainly determined by calculation of \mathbf{u} in comparison with the calculation of \mathbf{r}_0 . For less than approximately 5000 voxels, the real time demand $T_{\text{exec}} < 1$ second is fulfilled.

Figure 1 shows two exemplary synthetic data sets. In 1000 test cases, the difference angles between at random defined and determined direction vectors \mathbf{u} were less than 10° (median 3°), and the difference vectors between defined and determined origin point vector \mathbf{r}_0 had euclidean distances less than 0.6 (median 0.5).

4 Discussion

With the implemented algorithm, the real time demand $T_{\text{exec}} < 1$ second for the estimation of the 4 DoF position of a needle-like instrument can be fulfilled for up to 5000 scattered voxels detected in an ultrasound image volume.

Ongoing work is dedicated towards further minimization of the execution time T_{exec} , e.g. by a multithreading implementation of the algorithm and the detection of noise voxels.

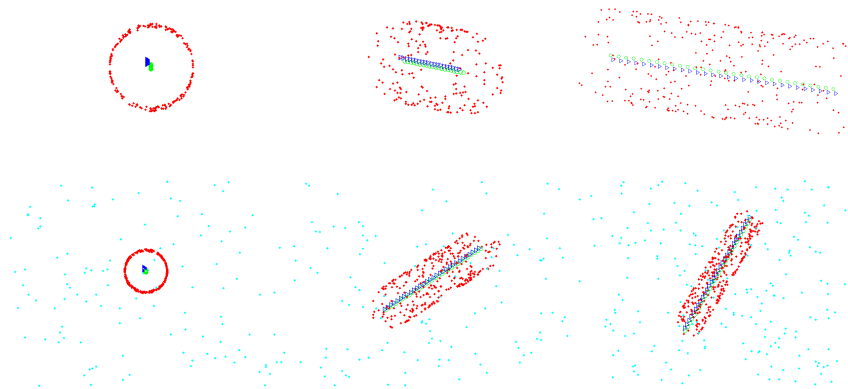


Fig. 1. Lines detected for synthetic tubes of length $L = 100$ and radius $R = 10$ with $N = 1000$ points. Upper row: Tube (red dots) with symmetry axis (green circles) and axis detected by the algorithm (blue triangles). View from above (left column) and two oblique views (center and right columns). Bottom row: the same as the upper row with additional noise data.

Table 1. Execution times.

Number of Points	Total Execution Time	Time for \mathbf{u}	Time for \mathbf{r}_0
N	$T_{\text{exec}}/\text{seconds}$	$T_{\mathbf{u}}/\text{seconds}$	$T_{\mathbf{r}_0}/\text{seconds}$
1000	0.117	0.116	0.000135
2000	0.238	0.238	0.000275
3000	0.431	0.431	0.000377
4000	0.679	0.679	0.000530
5000	1.021	1.020	0.000619
6000	1.380	1.375	0.000823
7000	1.816	1.808	0.000895
8000	2.325	2.311	0.001006
9000	2.844	2.842	0.001190
10000	3.541	3.539	0.001304

References

1. Brenner DJ. Radiation biology in brachytherapy. *J Surg Oncol.* 1997;65.
2. Hartmann P, Baumhauer M, Rassweiler J, et al. Automatic needle segmentation in 3D ultrasound data using a hough transform approach. In: *Proc BVM*; 2009. p. 341–5.
3. Hough P, inventor; P.V.C., assignee. Method and Means for Recognizing Complex Patterns, U.S.patent 3069654. U.S.patent 3069654, 1962; 1962.
4. Illingworth J, Kittler J. A survey of the hough transform. *Computer Vis Graph Image Process.* 1988;44:87–116.

High-Density Object Removal from Projection Images using Low-Frequency-Based Object Masking

Chris Schwemmer, Marcus Prümmer, Volker Daum, Joachim Hornegger

Pattern Recognition Lab, University of Erlangen-Nuremberg
`chris.schwemmer@informatik.stud.uni-erlangen.de`

Abstract. High-density objects, like catheters, pacemakers or even contrast agent-filled vessels, cause characteristic streak artifacts in computed tomography (CT). Similar to metal artifacts, these streaks can be reduced by removing the dense object using segmentation and interpolation. First, we compare state-of-the-art interpolation methods like linear, spline and higher-order methods to the Healing Brush technique. Second, a new method is presented, that extracts a low-frequency model of the dense object and restores the decomposed X-ray intensity of the remaining tissue. This method is henceforth called Subtract-and-Shift. Compared to standard interpolation methods, it retains the measured structure that is superimposed and dominated by the dense object. The extracted structure is then used to replace the segmented pixel intensities of the object. The introduced method is compared to state-of-the-art interpolation methods using in-vivo data. First preliminary results show that Subtract-and-Shift can be superior to these interpolation methods.

1 Introduction

Metal artifact reduction has been an active field of research since the beginnings of CT, especially in an interventional environment using C-arm CT. 3-D images are reconstructed from the projection images using filtered back-projection methods. These methods are sensitive to strong edges, like dense objects in a projection, due to a high-pass filtering of the measured projection data. Many of the clinical applications in C-arm CT have to deal with motion of organs like the heart or the lung. Dense moving objects, like a catheter in a heart, result in highly inconsistent data. A common solution to reduce resulting streak artifacts is to detect and replace this data with interpolated data. However, this is a challenging task, since important anatomical structure is often overlaid by the dense object and thus the underlying anatomy has to be guessed during interpolation to reduce artificial errors. Although many data interpolation methods exist, metal artifact reduction is usually done by linear interpolation [1]. Since, to the best of our knowledge, no quantitative comparison of interpolation methods for object removal from projection images exists, we are currently evaluating common interpolation methods as described in the next section. However, the

dense object does not always absorb the energy totally. A meaningful measurement can still be provided. In such a case, the anatomy of tissue is overlaid by the dense object. A catheter, for example, can be assumed to have a tubular structure. The idea of our new approach is to extract the underlying tissue measurement from the object intensity in the projection image. The shape of the dense object is modeled as a low-frequency bias value and subtracted from the measured intensity value. The remaining tissue intensity can provide meaningful structure for an improved data interpolation. This method is called Subtract-and-Shift and makes use of this remaining structure. We present first results of an in-vitro evaluation of standard interpolation methods in Section 3 and show that linear interpolation outperforms the other methods. Then we present first in-vivo results of Subtract-and-Shift and compare it to the three best performing standard interpolation methods, showing that structure is indeed retained by our method, whereas it is lost by using the other methods.

2 Materials and Methods

Line-wise linear interpolation, line-wise polynomial interpolation (degree 4), line-wise cubic b-spline interpolation (70 control points), defect pixel interpolation in frequency space [2] (2000 iterations) and the “Healing Brush” algorithm [3] have been evaluated and compared to each other using in-vitro data.

For the evaluation of interpolation results, we used the root-mean-square-deviation (RMSD) and the standard deviation of the difference image (SDDI). To do this, a catheter phantom was digitally inserted into a projection image and the interpolated image compared to the original, catheter-free image. Both low- and high-contrast phantoms were used to evaluate the interpolation performance.

Our new algorithm works as follows: (1) Blur a copy of the input image using a Gaussian window and subtract it from the input image at the pixels identified as belonging to the object to be removed. (2) For each image line, calculate the offset between all pixels left of the object and the object pixels, and the offset between all pixels right of the object and the object pixels. Then shift the intensity values of pixels belonging to the object, linearly interpolating between the two offsets.

We argue that the object to be removed introduces a low-frequency, model-based intensity bias. Although it follows from the attenuation law that a linear shift in *logarithmic* image space is introduced by an additional object in the X-ray beam, we hypothesize this can be treated as a linear shift in image space as an approximation. By subtracting the low-pass filtered version of the object area from the input, the low-frequency bias is removed, only retaining high-frequency content, i.e. structure that was still visible through the object, albeit centered around intensity value 0. The latter is corrected for by shifting the high-frequency content to the level of surrounding pixels’ intensities. Since important structure (e.g. bone, vessels, etc.) might pass through the affected area, this shift is adapted to both the left and right neighborhood by linearly interpolating between both offsets. As the preferred direction of filtering during

Table 1. Results of phantom interpolation (lc = low-contrast, hc = high-contrast).

Interpolation Method	RMSD (lc)	RMSD (hc)	SDDI (lc)	SDDI (hc)
(Catheter image)	16.5821	35.0539	16.5809	35.0387
Linear	13.0513	15.8172	13.0504	15.8148
Polynomial	37.8894	37.0169	37.8649	36.9951
Spline	15.4711	17.5983	15.4688	17.5943
Defect Pixel	17.2252	17.9420	17.2216	17.9374
Healing Brush	29.9780	30.6537	29.9699	30.6433

back-projection is horizontal, we use a line-wise approach to this offset correction. Figure 1 shows two in-vivo images used to evaluate this algorithm and compare it to common interpolation methods.

3 Results

Table 1 shows the interpolation results of the phantom images. The first line shows the values for RMSD and SDDI of catheter image vs. catheter-free (original) image to give an upper bound on these values.

Figures 2 and 3 show the results of applying Subtract-and-Shift to the two example pictures, while Fig. 4(a) shows an intensity plot of the marked area (Fig. 2(a)), visualizing the two steps of the algorithm. Finally, Figure 4(b) shows an intensity plot of the right part of the catheter (same image line as in Fig. 4(a)), including the result of our algorithm, as well as the result of linear, spline and defect pixel interpolation.

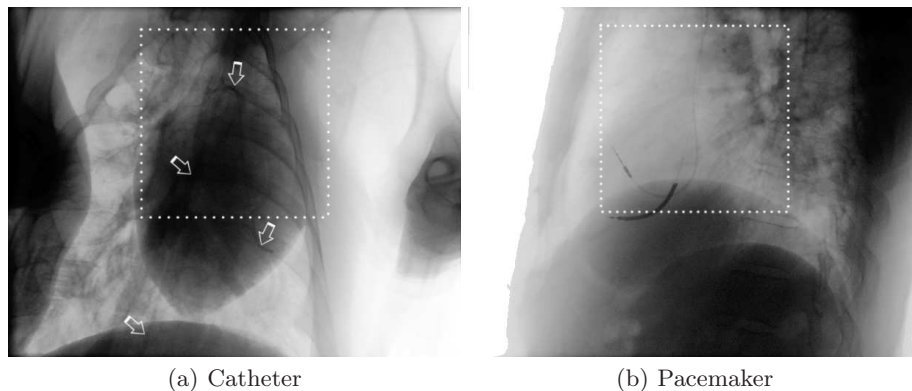


Fig. 1. In-vivo images used to evaluate Subtract-and-Shift. The dotted frame is the area shown in the images in the next section. In image (a), arrows point towards the catheter.

4 Discussion

From the absolute values of the results in Tab. 1, it can be seen that the interpolation results do not seem to be influenced much by the contrast of the phantom. It can also be seen that linear, spline and defect pixel interpolation perform best (in that order). While certainly producing visually pleasing results in photo manipulation, the Healing Brush introduced bigger differences to the original

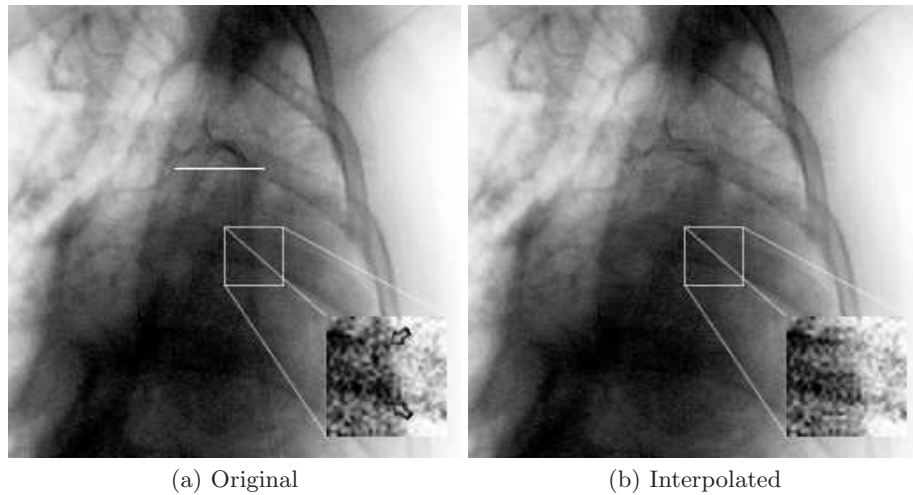


Fig. 2. Original and processed part from the catheter image. Line in (a) denotes area used for Fig. 4. Zoomed part is contrast enhanced to better show details. Arrows in zoomed part of (a) show points where catheter crosses the rib.

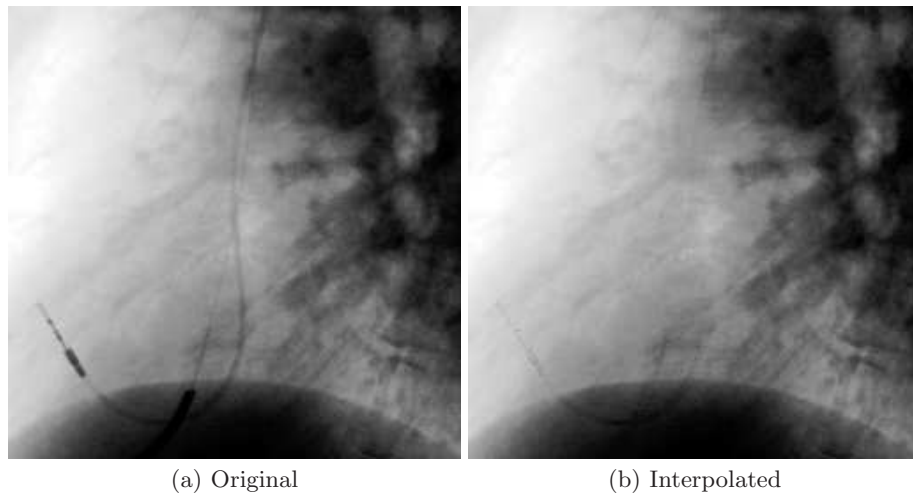


Fig. 3. Original and processed part from the pacemaker image.

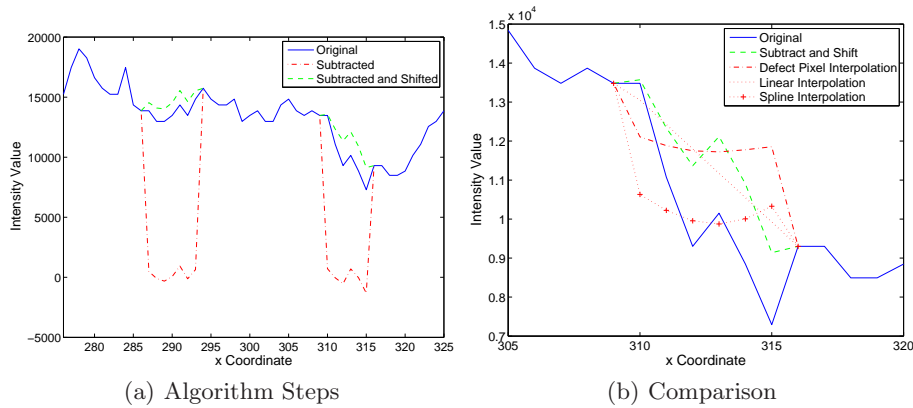
Fig. 4. Intensity plots of interpolation results.

image than the catheter phantom itself in the low-contrast test. But since this algorithm is by design highly dependent on the choice of “source material” to control the diffusion (c.f. [3]), further research on whether its performance in this environment can be improved is needed.

Figure 2 clearly shows that Subtract-and-Shift is able to reconstruct the information still contained inside the area of the catheter, e.g. parts of the ribs. As Fig. 4(b) confirms, the edge of the rib is reproduced (central spike), whereas the other interpolation methods do not reconstruct structure that corresponds to the underlying anatomy. The same can be observed for other parts of the image, where the catheter crosses tissue boundaries. Figure 3 show an equally good performance in the upper part of the pacemaker wire, albeit introducing artifacts in the bottom part.

For the future, we intend to produce ground-truth data from a physical phantom, to be able to compare our method to the others using RMSD or SDDI. Additionally, we intend to investigate the influence of different interpolation directions (as opposed to strictly horizontal) on the reconstruction result.

Acknowledgement. We would like to thank Rebecca Fahrig, Department of Radiology, Stanford University, for providing the projection images.

References

1. Kalender WA, Hebel R, Ebersberger J. Reduction of ct artifacts caused by metallic implants. *Radiology*. 1987;164:576–7.
2. Aach T, Metzler V. Defect interpolation in digital radiography: how object-oriented transform coding helps. *Proc SPIE*. 2001;4322:824–35.
3. Georgiev T. Image reconstruction invariant to relighting. In: *Proc Eurographics*. Dublin; 2005. p. 61–4.

Monte-Carlo-Based Scatter Correction for Quantitative SPECT Reconstruction Realization and Evaluation

Rolf Bippus¹, Andreas Goedicke¹, Henrik Botterweck²

¹Philips Research Laboratories, Aachen

²Fachhochschule Lübeck, Bildgebende Verfahren in der Medizintechnik

rolf.bippus@philips.com

Abstract. Quantitative SPECT as well as simultaneous acquisition of multiple isotopes with SPECT in the clinical field, although clinically interesting, are still limited by reconstruction artifacts and computing power. As a considerable step in this direction, we have implemented an efficient reconstructor with variance reduced Monte-Carlo-simulation in the forward and/or backward projection of an OS-EM iteration. Apart from a quantitatively accurate scatter estimation the integrated MC simulation allows us to include all effects that are relevant for the multiple isotope task. The reconstruction problem is explicitly formulated as a combined maximum-likelihood estimation for all unknowns in one step and implemented efficiently on the basis of the OS-EM algorithm using the Effective Scatter Source approach. The algorithm has been evaluated on simulated as well as measured phantom data. Two isotopes (Tc99m, Tl201) were tested on the voxelized NCAT phantom and evaluated quantitatively. In addition we performed phantom measurements to evaluate the method on a Philips CardioMDTM with a VantageTM line source for attenuation correction. A clear advantage of the proposed approach is its robustness and generalizability. It is currently being evaluated in clinical applications like simultaneous dual isotope cardiac imaging.

1 Introduction

Absolute quantification as well as simultaneous acquisition of multiple isotopes are receiving increasing interest in clinical applications (e.g. [1], [2]). However patient scatter, down-scatter plus collimator effects such as penetration and Pb-fluorescence result in data contamination.

Basically one can distinguish between energy and/or spatial distribution based scatter estimation and correction from reconstruction based methods [3]. The OS-EM (Ordered Subset Expectation Maximization) approach we follow with our implementation falls into the second category. Scatter is estimated based on the current activity estimate in the forward projection step. Using the approximative scheme of effective scatter source (ESS) Estimation, originally published by Frey et al. [4], scatter estimation is effectively separated from other effects in the algorithm. The original approach uses precomputed scatter kernels to

obtain the ESS via convolution. In contrast we apply variance reduced Monte-Carlo simulation to estimate the ESS. An attenuation map is used as a physical model of the patient.

Using one of the reconstruction based scatter estimation methods one can differentiate sequential and simultaneous reconstruction as depicted in Fig. 1. We follow the scheme of simultaneous reconstruction (Fig. 1(b)) where the cross contamination estimation is integrated into the forward projector and all estimates isotope activities are forward projected into all energy windows and updated in each sub-iteration of the OS-EM algorithm.

2 Methods

The full update equations for the ordered subset EM algorithm for all isotope activities (typically 1 or 2) $f_{\epsilon,i}$ (isotope i in voxel ϵ) on all available data are given by equations (1) and (2).

The forward projection step computes (1) the contributions of isotopes i to all projection pixels p in each energy-window e , based on the recent activity estimate $f_{\epsilon,i}^{(k)}$, where $H_{p,e}^{\epsilon,i}$ is the normalized system operator. In the backprojection step (2), the correction factors are obtained from the ratio of the observed and the estimated projection values.

$$\nu^{(k)}(p, e) = \sum_i \nu_i^{(k)}(p, e) = \sum_i \sum_{\epsilon} f_{\epsilon,i}^{(k)} H_{p,e}^{\epsilon,i} \quad , \quad \sum_{p,e} H_{p,e}^{\epsilon,i} = 1 \quad (1)$$

$$f_{\epsilon,i}^{(k+1)} = f_{\epsilon,i}^{(k)} \frac{1}{n_{\epsilon,i}} \sum_p \sum_e H_{p,e}^{\epsilon,i} \frac{\nu(p, e)}{\nu^{(k)}(p, e)} \quad , \quad n_{\epsilon,i} = \sum_{p,e} H_{p,e}^{\epsilon,i} \quad (2)$$

2.1 Approximations / Implementation

Using brute force counting statistics with MC-simulation to obtain the $\nu^{(k)}(p, e)$ in equation (1) is not feasible. Therefore we applied the principle of convolution based forced detection.

Simulating the photon tracks, at each interaction of the photons with the patient the effective cross section is calculated for the photon being redirected towards the detector for a set of intermediary energy intervals. These are accumulated in 3D and are treated as secondary emission distributions, the ESS. These are then projected onto all camera positions and energy windows, taking

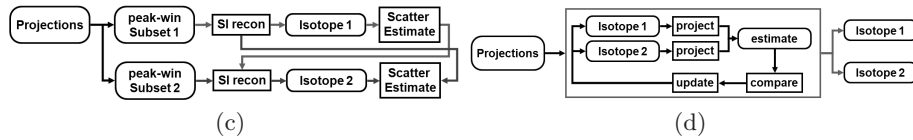


Fig. 1. Sequential (a) vs. simultaneous (b) dual isotope reconstruction scheme.

into account attenuation and collimator response functions (CRF). The CRFs, simulated beforehand and stored, are taking into account all major physical degradation effects. The principle is illustrated in Fig. 2.

According to the Dual-Matrix approach, the backprojector is approximated by the reversed PSFs and the attenuation ($B_{p,e}^\epsilon$) as well as the relative weights $\omega_i(p, e)$ of the isotope i contributing to pixel (p, e) in the forward projection, thus ignoring the spatial distribution of the scatter.

$$f_{\epsilon,i}^{(k+1)} = f_{\epsilon,i}^{(k)} \frac{1}{n_{\epsilon,i}} \sum_p \sum_e B_{p,e}^\epsilon \omega_i^{(k)}(p, e) \frac{\nu(p, e)}{\nu^{(k)}(p, e)} \quad (3)$$

2.2 Activity Estimation

Quantitative results are obtained by defining volumes of interest (VOI) and calculating the mean activity within the VOI. We evaluated quantitative activities after 30 to 50 full iterations of the OS-EM algorithm. Corresponding images are not suited for visual inspection or VOI delineation due to high noise. Thus images for visual inspection or delineation use 4 full iterations of the OS-EM algorithm. For the simulated data we defined the VOIs directly from true activity images.

3 Results

3.1 NCAT Phantom Results

The voxelized NCAT phantom was simulated such that the activities represent a patient dose of 25 mCi $\text{Tc}^{99\text{m}}$ and 3.5 mCi of Tl^{201} . In the Tl^{201} distribution a heart lesion was included, which was not present in the $\text{Tc}^{99\text{m}}$ image. The lesion activity was chosen 15 % of the normal myocardial Tl^{201} activity.

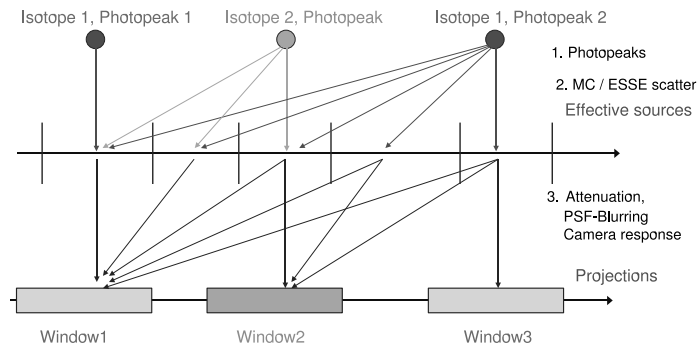


Fig. 2. Principle of down-scatter estimation in the forward projection step.

For quantitative comparison of true and reconstructed activity, we measured absolute activity in the 17 segments of a standard 17 segment polar plot representation. Heart re-orientation and segmentation were performed on the true activity image and kept constant for the reconstructed image. As before the values represent absolute activity estimates averaged over each segment of the heart.

3.2 Phantom Measurement Results

Phantom measurements were performed using a Data Spectrum Cooperation Anthropomorphic Torso PhantomTM with a fillable heart insert. The phantom was filled according to an estimated biodistribution resulting from 90 MBq Tl^{201} and 650 MBq Tc^{99m} administered to an adult. The heart contained 2 lesions of 5.6 ml (lesion 1) and 12 ml (lesion 2), each filled with Tc^{99m} only, which gives no Tl^{201} and approx. 65% Tc^{99m} (due to the walls of the lesions) in the lesions. Resulting images and polar plots are shown in Fig. 4.

4 Conclusion

Quantitative MC reconstruction for SPECT and multiple isotopes shows to be feasible. The more exact modeling of scatter within the object and collimator results in reduced image artifacts. The computational effort is reasonable, especially compared to other state of the art scatter corrections like ESSE.

An advantage of the proposed approach is its robustness and generalizability: Fast adaptation to new isotope combinations is possible. Next steps will include the evaluation of its potential in clinical application of dual isotope cardiac imaging.

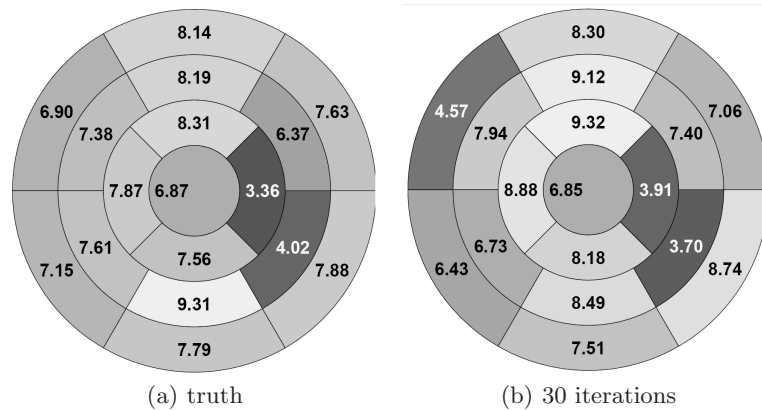
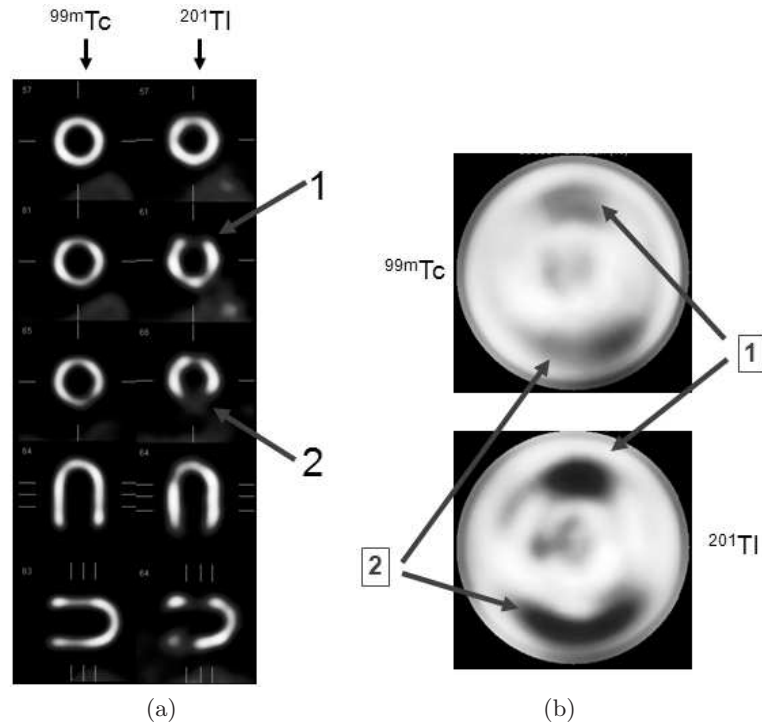


Fig. 3. NCAT, Tc^{99m} - Tl^{201} SDI, Tl^{201} mean activity in kBq/ml in each of 17 segments of standard polar plot.

Fig. 4. Phantom measurements on the CardioMD. Cardiac SDI filled with Tl^{201} and Tc^{99m} and two lesions filled with Tc^{99m} only.



References

1. Berman D, Kang X, Tamarappoo B, et al. stress thallium-201/rest technetium- 99m sequential dual isotope high-speed myocardial perfusion imaging. *JACC Cardiovasc Imaging*. 2009;2(3):273–82.
2. Sgouros G. Dosimetry of internal emitters. *J Nucl Med*. 2005;46(1):273–82.
3. King J M A ad Glick, Pretorius PH, Wells RG, et al. Attenuation, Scatter and Spatial Resolution Compensation in SPECT. In: Wernick M, Aarsvold JS, editors. *Emission Tomography. The Fundamentals of PET and SPECT*. Academic Press; 2004.
4. Frey EC, Tsui BMW. A new method for modelling the spatially variant, object-dependent scatter response function in SPECT. In: *Proc IEEE Nucl Sci Symp*; 1996. p. 1082–6.

Data Reduction for Supervised Learning in Medical Image Analysis

Introduction of a Measurement Decision Tree

Armin Stoll¹, Andrea Fränze¹, Rolf Bendl^{1,2}

¹Medical Physics in Radiation Oncology, DKFZ Heidelberg

²Medical Informatics, Heilbronn University

a.stoll@dkfz-heidelberg.de

Abstract. Segmentation of organs at risk for radiation therapy planning is a time-consuming task. Knowledge-based segmentation is in the focus of research to apply semi-automatic segmentation without user-interaction. It turns out that only integrating general knowledge into a knowledge base is not appropriate to deal with complex anatomical conditions among humans. Supervised learning is therefore investigated to provide individually given knowledge and adapt general knowledge to the individual case. Since deriving multi-dimensional feature spaces in medical images can cause huge amount of redundant data, a measurement decision tree is presented and its data reduction performance evaluated.

1 Introduction

Segmentation of organs at risk (OAR) for radiation therapy planning is a time-consuming and complex task. It is often done manually or semi-automatic. Volumes of interest (VOI) and their spacial location play a major role in radiation therapy planning. It enables to quantitatively simulate radiation therapy under certain therapy conditions. New developments in radiation therapy planning brought new requirements to the segmentation task on medical images. Automatic segmentation would allow to repeat treatment planning more often since this time-consuming task is still a limiting factor. Many semi-automatic segmentation algorithms are already developed. One drawback is that complex configuration can remain necessary to achieve expected segmentation results. As a consequence, semi-automatic segmentation also turns into a knowledge intensive task. Knowledge-based systems are in the focus of research to provide relevant knowledge. It is intended to trigger semi-automatic segmentation without user interactions [1]. But, it turns out that only incorporating general knowledge is not sufficient to deal with complex anatomical conditions among humans. Pattern recognition methods are therefore investigated to acquire individually given knowledge. As a result, semi-automatic segmentation algorithms are triggered by individually adapted knowledge.

This contribution introduces a measurement decision tree (MDT) for supervised learning purposes. It is intended to reduce the number of measurements for the training of classifiers on large scale medical image data.

2 Material and Methods

Knowledge-based segmentation [2, 3, 4] in radiation therapy planning is mainly investigated for computed tomography (CT) since it is the principal foundation for physical radiation dose simulation. CT images show X-ray attenuation coefficients in Hounsfield Units [HU] within a tiny volume – a voxel. To classify single transversal slices into body regions, different classifiers are already investigated [1, 5]. Further investigations, now focusing on multi-dimensional image processing and voxel classification, easily exceed the amount of training data by a factor of 2 to 10. This is due to statistical features (measurements) derived from small subregion for each voxel. For instance, minimum, maximum, median, variance and Haralick features [6]. In the same context, labels are derived from already existing manual segmentation (fig. 1). The amount of training data can easily turn into main memory requirements exceeding the 3 GB threshold on 32 bit computer systems. It is observed that among all measurements many are redundant represented (e.g. a lot of background measurements are fixed to -1024 HU and therefore highly over-expressed). Cross-validation on redundant data becomes very inefficient and turns into superfluous computational efforts.

3 Results

Decision trees are well known from the context of machine learning [2]. Unlike other classifiers, models produced by decision trees are easy to interpret. The reasoning process is transparent and straightforward from the root node to its leaf nodes. In similar prospect a MDT is proposed to arrange measurements m and its labels l in a tree representation to reduce redundancies.

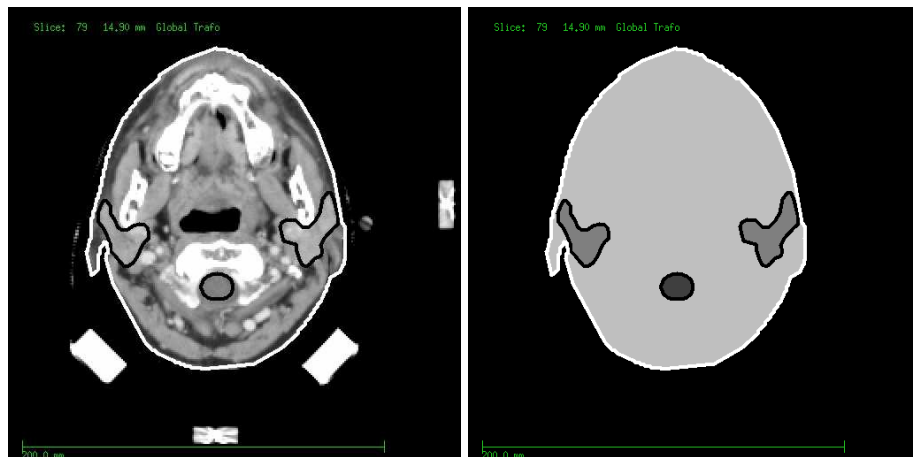


Fig. 1. Training data. Left: manually segmented contours of a head&neck case. Right: derived labels per voxel (background, human anatomy, spinal cord and parotids).

3.1 Measurement Decision Tree

Each measurement m consists of several features x_i . Each feature is represented on a separate level in the MDT (fig. 2). Two types of nodes are known: feature nodes and label nodes. A feature node (FN) consists of value x_i (e.g. a numerical value, character or text) and a list of subsequent feature nodes with values x_{i+1} . A label node (LN) is always a leaf node and contains a label l and a counter variable c . In Python syntax the implementation is the following:

```
class FeatureNode:
    def __init__(self, value, children=[]):
        self.value = value
        self.children = children
    def __eq__(self, value):
        return self.value == value

class LabelNode:
    def __init__(self, label, counter=1):
        self.label = label
        self.counter = counter
    def __eq__(self, label):
        return self.label == label
```

Training To build a MDT, a root node (RN) has to be created first. The first feature x_1 of the first measurement m_1 is attached to the RN as FN. The second feature x_2 is attached to the previous FN. This process is continued until every feature of the first measurement is integrated as the first branch in the MDT. Finally, a LN is attached to the last FN and its counter variable is initialized with value 1. Insertion of the second measurement m_2 begins at the RN again.

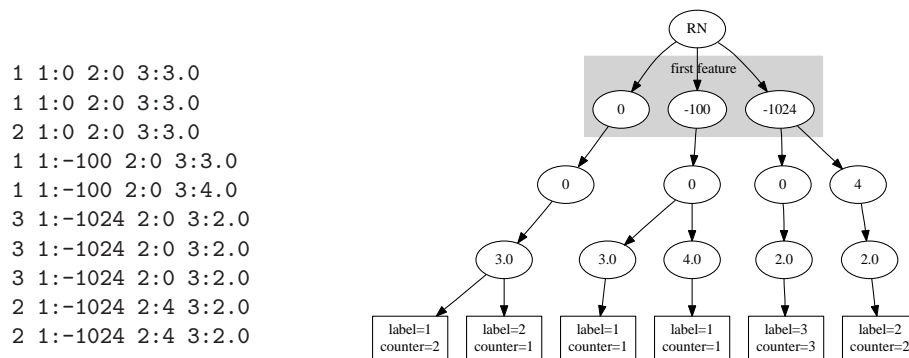


Fig. 2. Example: left some particular training data (format: label, feature1:value1, feature2:value2 etc.) and right the resulting MDT.

If the first feature is already present, no new node is created. If not, then a new FN is appended to the current node. This work flow is true for every feature x_i . The counter variable of a LN is either incremented by one if a corresponding LN already exists or a new one created. A simple example is shown in fig. 2. The implementation of the *insert* function is given as the following:

```
def insert(measurement, label):
    node = root
    for m in measurement:
        if m in node.children:
            i = node.children.index(m)
            node = node.children[i]
        else:
            node.children.append(Node(m, []))
            i = node.children.index(m)
            node = node.children[i]

    if label in node.children:
        i = node.children.index(label)
        node.children[i].value += 1
    else:
        node.children.append(LeafNode(label, 1))
```

As an advantage, no information get lost when measurements are represented as MDT. If more than one label is assigned to a measurement, different decision strategies become possible. For instance, conserving the measurement for every label, only for the label with maximum occurrence, for all labels that exceed a certain threshold or that one that gains absolute majority.

This MDT was applied to reduce the measurement space for five head&neck cases. Six features were derived within three-dimensional CT data for every voxel. The total number of measurements per case is depicted in tab. 1. The training data is reduced by integrating each measurement into a MDT and transforming it back into a linear sequence. Each label is now non-redundant represented by exactly one measurement. The total number of remaining measurements and its reduction is also shown in tab. 1. Finally, all reduced training sets are merged into one single training set that comprises all cases. The total number of remaining measurements is $6\,285 \cdot 10^3$. As a result, measurements from different classes are now more balanced distributed and redundancies strongly reduced.

4 Discussion

The proposed MDT allows to reduce redundant measurements on large scale medical image data to $\approx 7\%$ of its original size. Measurements are represented exactly once in the feature space. Adverse side effects on single classifiers (e.g. k-nearest neighbours classifier) have to be individually evaluated. The main benefit

Table 1. Evaluation: measurement reduction performance by insertion into the MDT.

Case	# Measurements	# Measurements remaining	Reduction [%]
1	$40 \cdot 10^6$	$1\,668 \cdot 10^3$	96
2	$21 \cdot 10^6$	$862 \cdot 10^3$	96
3	$28 \cdot 10^6$	$1\,151 \cdot 10^3$	96
4	$25 \cdot 10^6$	$1\,018 \cdot 10^3$	96
5	$41 \cdot 10^6$	$6\,381 \cdot 10^3$	84
Mean:	$31 \cdot 10^6$	$2\,216 \cdot 10^3$	93

results in the reduction of computational effort for supervised learning and cross-validation since the chance of testing redundant data reduces tremendously. As a result, individual knowledge acquisition on voxel scale becomes possible for knowledge-based segmentation purposes [2, 4].

References

1. Stoll A, Bendl R. Wissensakquisition mit methoden der mustererkennung zur wissensbasierten segmentierung von risikoorganen in CT-Bilddaten. Proc BVM. 2008; p. 41–5.
2. Schmidt G, Kietzmann M, Kim J, et al. Cognition network technology for fully automatic 3D segmentation of lymph nodes in CT data. In: Dössel O, Schlegel WC, editors. World Congr Med Phys Biomed Eng. vol. 25/IV; 2009. p. 1365–7.
3. Foruzan AH, Zoroofi RA, Hori M, et al. A knowledge-based technique for liver segmentation in CT data. Comput Med Imaging Graph. 2009;33(8):567–87.
4. Klinder T, Wolz R, Lorenz C, et al. Spine segmentation using articulated shape models. In: Proc MICCAI; 2008. p. 227–34.
5. Fränze A, Stoll A, Bendl R. Ermittlung einer kranial-kaudalen Korrespondenz in MR-Aufnahmen. Proc BVM. 2009; p. 232–6.
6. Medical image analysis of 3D CT images based on extension of Haralick texture features. Comput Med Imaging Graph. 2008;32(6):513–20.

Accelerated C-Arm Reconstruction by Out-of-Projection Prediction

Hannes G. Hofmann, Benjamin Keck, Joachim Hornegger

Pattern Recognition Lab, University Erlangen-Nuremberg
hannes.hofmann@informatik.uni-erlangen.de

Abstract. 3D image reconstruction from C-arm projections is a computationally demanding task. However, for interventional procedures a fast reconstruction is desirable. We present a method to reduce the number of actually back-projected voxels and ultimately the processing time by 20–30% without affecting image quality. The proposed method detects and skips subvolumes that are not visible in the current view. It works with projection matrices and is thus capable of handling arbitrary geometries. It can easily be incorporated into existing algorithms and is also suitable for the back-projection step of iterative algorithms.

1 Introduction

Clinical applications – in particular interventional procedures – require fast 3D reconstruction of tomographic data. Therefore, means to reduce the reconstruction time are a strong research topic. Currently, the most wide-spread algorithms in clinical X-ray CT reconstruction belong to the class of filtered back-projection (FBP), e.g. the FDK method [1] for cone-beam data. Another class of reconstruction algorithms is based on iterative techniques which require multiple alternating back- and forward-projections. The latter can incorporate various corrections and provide superior image quality in certain cases like sparse or irregular data [2]. The fact that their complexity is a multiple of the FDK’s is a reason why this class of reconstruction algorithms is less commonly used in clinical systems. However, both classes of reconstruction algorithms benefit from an acceleration of the back-projection step.

Several groups investigate acceleration using specialized hardware (e.g. FPGAs [3, 4, 5]), or general purpose high-performance hardware (Cell processor [6, 7], GPUs [8, 9, 10], Larrabee [11]).

Yu et al. [12] proposed to reconstruct only cylinder- or sphere-shaped region of interest – approximating the Field-of-View (FOV) of the scan – to reduce the number of back-projected voxels. Their method, however, requires a priori knowledge about the trajectory and offline computation of the FOV before reconstruction.

Yu et al. further suggested data partitioning which is similar to the subvolumes approach used e.g. by Scherl et al. [6] and Kachelriess et al. [7]. They used subvolumes to fit the problem into the 256 KB-sized local stores of the Cell

Algorithm 3 Original back-projection step, called once for each projection image I_n .

```

for all voxels  $\underline{x}$  do
  Project voxel onto detector plane
  Check if point lies on detector
  Fetch projection values
  Bilinear interpolation
  Update  $f_{\text{FDK}}(\underline{x})$ 
end for

```

processor. They also proposed that subvolumes improve caching on CPUs resulting in better performance. However, they did not actually skip subvolumes as proposed in this paper.

Our proposed method adapts dynamically to the FOV of each projection image during back-projection. It uses projection matrices from the scanner and requires no prior knowledge about the acquisition trajectory. It is easy to integrate into existing reconstruction algorithms and does not affect image quality.

2 Materials and Methods

Alg. 3 shows pseudo-code of the back-projection step of the FDK method. It is called once for each projection image I_n and updates every voxel of the reconstructed volume.

In cone-beam CT, the source position and the corners of a projection image define a pyramid, the current FOV. Areas of the reconstructed volume that are outside of the FOV are not visible on the detector. Hence, the current view cannot contribute information about their content.

Alg. 4 shows the modified back-projection step. Before a subvolume is back-projected a check is performed to decide whether it is inside the current FOV. Therefore, its corner voxels are projected to figure out its shadow. If the intersection of the shadow with the detector has a positive area it is back-projected as usual. Otherwise, the whole subvolume is skipped and processing continues with the next one. The computation of a subvolume's shadow involves 3D/2D-projections of its 8 corners, computation of the minimum and maximum of the resulting 2D coordinates, clamping of the coordinates and computation of the shadow's area. On the other hand, for every voxel contained in a skipped subvolume a 3D/2D-projection, the coordinates check, $4 \times$ pixel access, bilinear interpolation and the voxel update are saved.

The proposed method skips only subvolumes that did not contribute to the current projection. Hence, there should be no difference in the result compared to the reference method. The reconstruction benchmark RABBITCT [13] was used to test our implementation for correctness and measure its performance.

Obviously, the performance of our new method depends on the acquisition geometry. To rule out the possibility of "cooperative data" the method was evaluated on two other clinical datasets, additional to the public RABBITCT

Algorithmus 4 Modified back-projection step, called once for each projection image I_n .

```

Partition volume  $f_{\text{FDK}}$  into subvolumes  $f_{\text{FDK},i}$ 
for all subvolumes  $f_{\text{FDK},i}$  do
  // Test if  $f_{\text{FDK},i}$  is visible
  Compute shadow of  $f_{\text{FDK},i}$  by projecting corners
  if  $\emptyset \neq \text{shadow} \cap I_n$  then
    Skip  $f_{\text{FDK},i}$ 
  else
    for all voxels  $\underline{x} \in f_{\text{FDK},i}$  do
      Project voxel onto detector plane
      Check if point lies on detector
      Fetch projection values
      Bilinear interpolation
      Update  $f_{\text{FDK}}(\underline{x})$ 
    end for
  end if
end for

```

dataset. All datasets were acquired with real C-arm systems using clinical protocols. Their sizes are shown in Tab. 1 where dataset A is the public RABBITCT dataset.

To show that the new method also improves the performance of highly optimized code it was incorporated into our vectorized and multi-threaded FDK implementation for CPUs (described in more detail in [14]).

The new method was tested on two multi-core systems. The first system featured two Core i7 (“Nehalem”) quad-core processors at 2.66 GHz (85.12 GFlops total) and was equipped with 12×1 GB of DDR3-1066 RAM. This system had a very high memory bandwidth. The other one had four hexa-core Xeons (X7460) at 2.66 GHz (255.35 GFlops total) and 32 GB of DDR2-1066 RAM. All tests were performed using 64-bit Linux and the Intel compilers in version 11.1.056. The “X7460” and “Nehalem” systems used in this report are preproduction systems. We expect production hardware to deliver similar performance levels.

3 Results

The reconstructed volume had a size of 512^3 voxels. Tab. 1 shows the results for both computer systems. The subvolume size was empirically chosen to be $128 \times 64 \times 4$ as this performed overall best. In the worst case, this subvolume size was about 5% slower than the best one. The large size in x -direction is justified by the fact that neighboring voxels in this direction are stored successively in memory and therefore supportive for the hardware prefetching mechanism.

Results for three different implementations are shown. The baseline is defined by our vectorized and multi-threaded implementation [14] (“base” in Tab. 1). Using subvolumes allows to optimize the cache usage of modern processors by

Table 1. Description of the three datasets. Further, reconstruction times (seconds). Results are shown for two systems (“Nehalem” and “X7460”) and three implementations (“base”, “cache” and “skip”). The volume size was 512^3 , the subvolume size $128 \times 64 \times 4$.

Dataset	Number of projections	Projection size	“Nehalem”			X7460		
			base	cache	skip	base	cache	skip
A	496	1240×960	84.71	80.28	61.38	60.53	42.89	31.96
B	543	1240×960	92.89	87.56	67.39	66.04	48.77	35.08
C	414	1024×1024	70.54	68.23	54.54	50.59	36.05	29.21

re-using voxel data before storing it back to main memory. Consequently, the memory bandwidth restriction is relaxed (“cache”). Finally, the results of the method proposed in this work (“skip”) are presented.

The optimized cache usage gives a speedup of at least $1.35 \times$ on the bandwidth-starving “X7460” system. On the “Nehalem” system, however, it shows only little effect (up to $1.06 \times$). This result was expected due to the higher memory bandwidth of the Nehalem architecture.

With the proposed method of skipping subvolumes the reconstruction time is further reduced to about 72–80% compared to the cache-optimized version on both systems (1.23 – $1.39 \times$). The effectiveness of our approach can be seen by looking on the numbers of a single dataset, e.g. A. Given the chosen size of $128 \times 64 \times 4$ voxels about 24% of the subvolumes are outside of the FOV. The reconstruction time is 75 and 76% of the “cache” implementation’s on the “X7460” and the “Nehalem” system, respectively.

4 Discussion

We have tried numerous different subvolume sizes and shapes and found out that the optimal solution depends on several factors. Large subvolumes – e.g. $512 \times 64 \times 8$ – provide better data locality and use less memory bandwidth if caches are large enough. However, this advantage is absorbed by the fact that they are less likely to be skipped since smaller subvolumes resemble the real FOV more accurately. While small subvolumes can have skip-ratios of up to 32% (for $16 \times 16 \times 4$ and dataset A) they come with the overhead of more shadow computations and with reduced savings per skipped subvolume.

On the other hand, subvolume shape has an impact on other performance factors. From a theoretical point, square-shaped subvolumes should be optimal as their mean shadow size over all viewing angles is minimal. But given current CPUs’ caching and prefetching algorithms it is a good idea to use subvolumes which are longer in the direction of contiguous memory access.

The savings of the proposed method depend on the acquisition geometry, namely the ratio between volume and FOV sizes. To show its relevance the method was tested using three different clinical datasets. The proposed method delivers a considerable speedup even to highly optimized implementations. Al-

though all datasets used here were circular trajectories the method is suitable for arbitrary trajectories. It requires no prior knowledge other than the projection matrices. Since it works on-the-fly it can easily be incorporated into existing algorithms.

References

1. Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. *J Opt Soc Am.* 1984;A1(6):612–9.
2. Kunze H, Härer W, Stierstorfer K. Iterative extended field of view reconstruction. *Proc SPIE.* 2007;6510:5X–1–8.
3. Xue X, Cheryauka A, Tubbs D. Acceleration of fluoro-CT reconstruction for a mobile C-Arm on GPU and FPGA hardware: A simulation study. *Proc SPIE.* 2006;6142:1494–501.
4. Churchill M. Hardware-accelerated cone-beam reconstruction on a mobile C-arm. *Proc SPIE.* 2007;6510:5S–1–8.
5. Heigl B, Kowarschik M. High-speed reconstruction for C-arm computed tomography. *Radiol Nucl Med.* 2007; p. 25–8.
6. Scherl H, Koerner M, Hofmann H, et al. Implementation of the FDK algorithm for cone-beam CT on the cell broadband engine architecture. *Proc SPIE.* 2007;6510:58–1–10.
7. Kachelrieß M, Knaup M, Bockenbach O. Hyperfast parallel-beam and cone-beam backprojection using the cell general purpose hardware. *Med Phys.* 2007;34(4):1474–86.
8. Mueller K, Yagel R. Rapid 3D cone-beam reconstruction with the algebraic reconstruction technique (ART) by utilizing texture mapping graphics hardware. *IEEE Nucl Sci Symp Conf Rec.* 1998;3:1552–9.
9. Xu F, Mueller K. Real-time 3D computed tomographic reconstruction using commodity graphics hardware. *Phys Med Biol.* 2007;52(12):3405–19.
10. Scherl H, Keck B, Kowarschik M, et al. Fast GPU-Based CT reconstruction using the common unified device architecture (CUDA). In: *IEEE Nucl Sci Symp Conf Rec.* vol. 6; 2007. p. 4464–6.
11. Hofmann HG, Keck B, Rohkohl C, et al. Towards C-arm CT reconstruction on Larrabee. In: *Proc Fully 3D Meeting and HPIR Workshop; 2009.* p. 1–4.
12. Yu R, Ning R, Chen B. High speed cone beam reconstruction on PC. *Proc SPIE.* 2001;4322:964–73.
13. Rohkohl C, Keck B, Hofmann HG, et al. RabbitCT: an open platform for benchmarking 3D cone-beam reconstruction algorithms. *Med Phys.* 2009;36(9):3940–4.
14. Hofmann HG, Keck B, Rohkohl C, et al. Putting 'p' in rabbitCT: fast CT reconstruction using a standardized benchmark. In: *PARS-Mitteilungen.* GI, Gesellschaft für Informatik e.V.; 2009. p. in press.

Ganganalyse mit Tiefendaten

Jochen Radmer¹, Jörg Krüger^{1,2}

¹Institut für Werkzeugmaschinen und Fabrikbetrieb, TU Berlin

²Fraunhofer Institut für Produktionsanlagen und Konstruktionstechnik
Forschungsgruppe Rehabilitationsrobotik (Fraunhofer IPK/TU Berlin)

radmer@iwf.tu-berlin.de

Kurzfassung. In diesem Beitrag werden die Arbeiten zur Umsetzung eines 3D-Bewegungsanalysesystems auf Basis einer Tiefenkamera vorgestellt, das den Anforderungen für den Routinebetrieb im klinischen Alltag gerecht werden soll. Wichtige betrachtete Kriterien sind dabei: einfache Integration in ein vorhandenes klinisches Umfeld, kein Rüstaufwand am Patienten und schnelle Durchführung der Messung sowie die Genauigkeit der Messergebnisse.

1 Einleitung

1.1 Einführung in das Problemfeld

Bei den Arbeiten im Bereich der Ganganalyse ist zwischen den klinischen Forschung und der klinischen Anwendung zu unterscheiden. Die klinische Forschung umfasst die Forschung zum Verständnis der Auswirkungen von Parametern auf eine Patientengruppe oder den Effekt einer Behandlung, wohingegen die klinische Anwendung sich mit der Umsetzung von klinisch anwendbaren Lösungen beschäftigt. Die Kriterien bei der klinischen Forschung und der klinischen Anwendung sind daher nicht zwingend identisch. Viele Arbeiten zur optischen Ganganalyse finden derzeit zu klinischen Forschung statt, nur wenige zu klinischen Anwendung.

In der klinischen Anwendung erfolgt derzeit die Ermittlung von Therapiefortschritten bei der Gangrehabilitation von Schlaganfallpatienten fast ausschließlich über qualitative Skalen und Scores. Das größte Problem dabei ist die vergleichsweise geringe Qualität der Messergebnisse, da diese inhärent subjektiv sind und zudem eine hohe Varianz aufweisen. Ursache dafür sind die Inter- sowie die Intra-Rater Variability. Aus beiden sich überlagernden Aspekten folgen signifikante Defizite hinsichtlich der Diagnose von Patienten-Bewegungsdefiziten, die die weitere Planung der motorischen Rehabehandlung für Schlaganfallpatienten unterstützt. In der klinischen Forschung werden gegenwärtig überwiegend Verfahren eingesetzt, die das Anbringen von Markern bzw. Befestigungsstrukturen am Patienten sowie eine geeignete Laborausstattung erfordern und die Gelenkbewegungen aus den Bewegungen der Marker schließen.

Diese Systeme haben hohe Anschaffungskosten und sind zudem durch das Platzieren der Marker bzw. der Befestigungsstrukturen sehr zeitaufwendig. Zusätzlich zu den Fehlern, die durch die Relativbewegung zwischen Weichteilen wie

der Haut und den Knochen induziert werden und dadurch zu einem maximalen Gelenkwinkelfehler von 7° führen können, bergen sie das Risiko, künstlich die Bewegungsabläufe zu beeinflussen und damit zu verfälschen [1]. Aus diesem Grund bilden markerbasierte Verfahren für die klinische Anwendung keine praktikable Basis. Aufgrund des Fehlens dieser systembedingten Eigenschaften von markerbasierten Ansätzen bilden markerfreie Ansätze für die klinische Anwendung die einzige Alternative. Einige der wenigen markerfreien Ansätze, die auf die klinische Anwendung abzielen, werden exemplarisch im folgenden Abs. 1.2 zum Stand der Forschung kurz vorgestellt.

In diesem Beitrag werden die Arbeiten zur Umsetzung eines 3D-Bewegungsanalysesystems vorgestellt, das den Anforderungen für den Routinebetrieb im klinischen Alltag gerecht werden soll. Wichtige betrachtete Kriterien sind dabei: einfache Integration in ein vorhandenes klinisches Umfeld, kein Rüstaufwand am Patienten und schnelle Durchführung der Messung sowie die Genauigkeit der Messergebnisse. Das System wird dabei zunächst für die Erfassung der Bewegungsabläufe eines Patienten auf einem Rehabilitationsroboter ausgelegt (siehe Abb. 1). Perspektivisch soll dieses System jedoch auch für die mobile Anwendung einsetzbar sein.

Abb. 1. Patient auf dem Gangrehabilitationsroboter *Haptic Walker*.



1.2 Stand der Forschung

Es existieren viele Ansätze zur Schätzung menschlicher Bewegungen, die oft auch Leistungsvermögen für die klinische Anwendung besitzen. Einen guten Überblick über die verschiedenen Ansätze bietet u. a. Poppe [2]. Sowohl Einzel- als auch Mehrkamarasysteme wurden bereits für die Messung biomechanischer Größen eingesetzt. 2D-Einzelkamarasysteme erfassen jedoch nur eine 2D-Projektion der dreidimensionalen menschlichen Bewegung und verlieren dadurch Information in der Tiefe, was nur bedingt über den Einsatz eines parametrisierten Modells kompensiert werden kann. Daher finden 2D-Einzelkameraverfahren nur für definierte Szenarien Anwendung, so dass meist Mehrkamarasysteme verwendet werden. Ein solches Mehrkamarasystem, das aus der klinischen Forschung in die

klinische Anwendung geführt werden soll, ist in der Arbeit von Mündermann et al. [3] beschrieben. Bei dem System handelt es sich um ein modellbasiertes System, das ebenso wie der gewählte Ansatz einen Abgleich zwischen Sensordaten und einem Modell im 3D durchführt. Als 3D-Representation wird dabei eine visuelle Hülle (visual hull) verwendet, die über Shape-from-Silhouette erzeugt wird. Für die Synthese der Körperhaltung wird sowohl ein Articulated Iterative Closest Point (ICP) als auch Simulated Annealing verwendet. Nach unserer Ansicht bieten Mehrkamerasysteme wie das von Mündermann et al. ein hohes Potential für die klinische Anwendung. Allerdings, wie auch von Mündermann et al. verfolgt, eher für die stationäre klinische Anwendung, da Mehrkamerasysteme einen relativ hohen Integrationsaufwand haben. Für kleine Messvolumina, die mit 3m x 1,5m x 2m für die Erfassung eines gesamten Gangzyklus ausreichen, bieten Tiefenkameras hier eine Alternative, die daher in dieser Arbeit betrachtet werden sollen. Eine Arbeit, die ebenfalls auf Tiefenkameradaten beruht, ist die von Jensen et al. [4]. Sie verwendet *PoseCut*, das sowohl Tiefen- als auch Intensitätsdaten einbezieht und die Körperhaltung über ein Stick-Modell mit Hilfe von Markov Random Fields synthetisiert. In der Arbeit wird eine auf einem Laufband gehende Person von der Seite betrachtet. Da es sich um ein zyklisches 2D-Stick-Modell handelt, das zuvor für das zu analysierende Gangmuster erzeugt werden muss, sind mit diesem Ansatz die pathologischen sowie die individuellen Ausprägungen eines Gangbildes nicht ausreichend erfassbar und damit der Ansatz für die klinische Forschung und Anwendung nicht geeignet.

2 Material und Methoden

Die Umsetzung des Systems basiert sensoruell auf einer Tiefenkamera, die aufgrund der geringeren Komplexität gegenüber einem Mehrkamerasystem der Forderung nach einer einfachen Integration entspricht. Momentan stellen Tiefenkameras jedoch noch Herausforderungen, wie der geringe Erfassungsbereich und die Datenqualität, die sich negativ auf die Anwendbarkeit auswirken. So weisen die Daten von Tiefenkameras ein relativ hohes Maß an Rauschen auf und besitzen Fehlerquellen, die bisher noch nicht vollständig kompensiert werden konnten. Die Hersteller sowie viele Forschungsgruppen nehmen sich jedoch dieser Thematik an und haben bereits leistungsfähige Verfahren vorgestellt, wie bspw. [5]. Daher wird davon ausgegangen, dass kommende Generationen von Tiefenkameras eine weit höhere Datenqualität aufweisen und damit deren Anwendbarkeit steigt.

Für die Umsetzung der Bewegungserfassung wird ein Analyse-durch-Synthese Ansatz verfolgt, der die Körperhaltung für jeden Frame einzeln schätzt. Der Einfluss vorheriger Schätzungen wird aufgrund der stark individuellen Ausprägung der Gangbewegung und der hinzukommenden Problematik der Generierung von Stichproben auf biomechanische Bewegungsbereiche des verwendeten Modells sowie die Initialisierung der Schätzung über die vorherige Schätzung beschränkt. Dazu wird ein Oberflächen-Mehrkörpermodell aus Ellipsenkegelstümpfen mit 24 Freiheitsgraden verwendet [6]. Um der Relevanz der Hüftbewegung in der Gangrehabilitation gerecht zu werden und eine möglichst geringere Selbstverdeckung

zu haben, wird der Proband aus frontaler Perspektive erfasst. Die Synthese erfolgt über eine Regression des Mehrkörpermodells in die Tiefendaten, die über die Minimierung einer Energiefunktion erfolgt, wie sie ähnlich in [7] vorgestellt wurde. Prinzipiell sind auch andere Energiefunktionen geeignet. Die Wahl der Energiefunktion fiel auf ICP aufgrund der einfachen Implementierung und Fusionierbarkeit verschiedener Daten. Die Initialisierung für den ersten Frame erfolgt momentan durch Benutzerinteraktion. Zur automatisierten Initialisierung sowie Parameterisierung der Modelle existieren jedoch leistungsfähige Arbeiten, wie das von Anguelov et al. [8]. Für die Synthese werden bereits segmentierte Daten verwendet. Die Segmentierung erfolgt dabei über einen Meanshift-basierten Ansatz, der einer Erweiterung des von von Bleiweiss et al. [9] vorgestellten Ansatzes entspricht. Dabei werden zusätzlich zu Tiefen- und Intensitätsdaten Bewegungsdaten genutzt, deren Einfluss variabel über ein Vertrauensmaß bestimmt wird. Für die Bewegungsdaten wird als Vertrauenswert die Amplitude einer Sobel-Kantenfilterung verwendet, der eine kantenerhaltende bilineare Filterung vorgeschaltet ist. Eine Erhöhung der Robustheit der Synthese wird durch die Verwendung von Deskriptoren (SURF) erreicht, die über aufeinander folgende Frames verfolgt werden, und es dadurch ermöglichen, die Rotationsbewegung des Beckens während eines Gangzyklus zusätzlich zur biomechanischen Schätzung über das Modell zu erfassen. Dazu werden die Deskriptoren als virtuelle Marker, solange existent, über mehrere Frames getrackt und dynamisch Stellen auf dem Modell rückwirkend über die Frames zugewiesen.

3 Ergebnisse

Für die Anwendung in der Ganganalyse sind die Knie- sowie die Hüftgelenkwinkelverläufe die beiden wichtigsten Größen bei der Gangrehabilitations-Therapie. Die Robustheit des Verfahrens soll daher anhand dieser Größen gezeigt werden. Bei den hier dargestellten Ergebnissen führte der Proband das Gangmuster *Gehen in der Ebene* mit einer Kadenz von 60 Schritten die Minute durch. In Abb. 2 auf der linken Seite ist die Beckenrotation in transversaler Ebene über die Zeit für

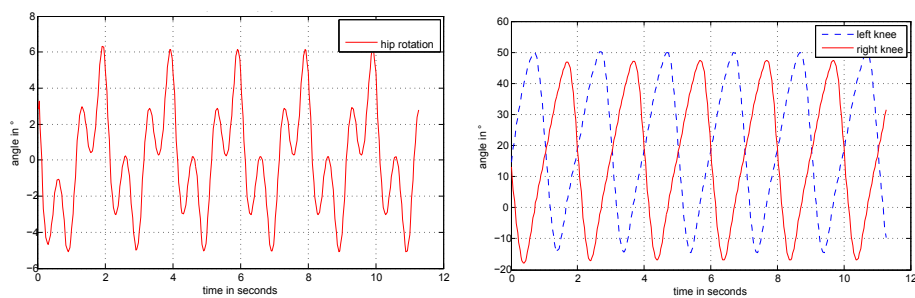


Abb. 2. Beckenrotation (links) und der Flexionswinkel des linken und rechten Kniegelenks (rechts).

mehrere Gangzyklen dargestellt. Im Vergleich zu Referenzdaten aus [10] weist die durch das markerfreie System erfasste Beckenrotation dieselben Charakteristiken auf. Auf der rechten Seite der Abb. 2 sind die Flexionswinkel des rechten und linken Knies zu sehen. Beider Flexionswinkel zeigen ein symmetrisches Verhalten und entsprechen abgesehen von einer kaum ausgeprägten Standphasenflexion den Referenzdaten.

4 Diskussion

In diesem Artikel wurden die Arbeiten zur Umsetzung eines tiefenkamerabasierten Systems zur markerfreien Ganganalyse für die klinische Anwendung vorgestellt. Dazu wurde zunächst als Anwendung der Gangrehabilitationsroboter *Haptic Walker* betrachtet. Das System war dabei in der Lage Gelenktrajektorien und -winkel markerlos zu erfassen, die in ihren Charakteristika ähnlich zu Referenzdaten aus der Literatur sind. Die Deformation der Knieflexionstrajektorie muss noch untersucht werden. Insbesondere in Bezug auf den zumindest teilweisen Einfluss des Gangrehabilitationsroboters auf das Gangbild. Eine Evaluation anhand eines markerbasierten Systems steht noch aus. Darüber hinaus ist eine Verschiebung der Knieflexionstrajektorien zwischen linkem und rechten Bein zu beobachten, der noch zu kompensieren ist.

Literaturverzeichnis

1. Leardini A, Chiari L, Della Croce U, et al. Human movement analysis using stereophotogrammetry: Part 3. Soft tissue artifact assessment and compensation. *Comput Methods Prog Biomed.* 2005;21(2):212–25.
2. Poppe RW. Vision-based human motion analysis: An overview. *Int J Comput Vis Image Underst.* 2007;108(1-2):4–18.
3. Mündermann L, Corazza S, Andriacchi T. The Evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications. *J Neuro Eng Rehabil.* 2006;3(1):1–11.
4. Jensen RR, Paulsen RR, Larsen R. Analysis of gait using a treadmill and a time-of-flight camera. In: *Proc DAGM Workshop Dyn3D; 2009.* p. 154–66.
5. Radmer J, Fuste PM, Schmidt H, et al. Incident light related distance error study and calibration of the PMD-range imaging camera. In: *Proc IEEE CVPR; 2008.* p. 1–6.
6. Balan AO, Sigal L, Black MJ, et al. Detailed human shape and pose from images. In: *Proc IEEE CVPR; 2007.*
7. Demirdjian D, Ko T, Darrell T. Constraining human body tracking. In: *Proc IEEE ICCV; 2003.* p. 1071.
8. Anguelov D, Srinivasan P, Koller D, et al. SCAPE: shape completion and animation of people. *ACM Trans Graph.* 2005;24(3):408–16.
9. Bleiweiss A, Werman M. Fusing time-of-flight depth and color for real-time segmentation and tracking. In: *Proc DAGM Workshop Dyn3D; 2009.* p. 58–69.
10. Rose J, Gamble JG. *Human Walking.* Williams & Wilkins; 1993.

Speckle Reduction for Automated Breast Ultrasound A Comparison of Three Filtering Methods

Fabian Zoehrer, Johann Drexl, Horst Hahn

Fraunhofer MEVIS, Universitaetsallee 29, Bremen, Germany
fabian.zoehrer@mevis.fraunhofer.de

Abstract. This paper compares filtering methods for automated breast ultrasound images. The Perona-Malik diffusion filter, the diffusion stick filter and a simple median filter are evaluated as preprocessing methods for segmentation and characterization of lesions in breast ultrasound. In particular, preservation of small relevant image features, edge enhancement characteristics, and computational costs are considered.

1 Introduction

Compared to other common modalities used in medical imaging, the strong signal to noise ratio of ultrasound is affecting its imaging quality to a large extent. Typical ultrasound images are often described as coarse-grained and diffuse due to multiplicative noise called speckle [1, 2]. Although the speckle pattern is characteristic for the imaged tissue, it makes image processing more complicated. Frequently, methods used for computer aided detection (CAD), segmentation, registration, feature extraction or classification of medical data require preprocessing steps for speckle reduction and/or edge enhancement. Among the commonly used methods are simple median filtering, methods based on wavelets [3], nonlinear diffusion [4, 5, 6], stick filter [7] and others [8]. As these methods differ largely in terms of filtering quality and computational costs, none of them can be said to be the dominant technique for all applications. The goal of this paper is to compare a set of filters for (automated) breast ultrasound.

2 Materials and Methods

The SomoVu automated breast ultrasound system from U-Systems Inc. (San Jose, CA, USA; EC Representative: Siemens, Erlangen, Germany) generates three-dimensional ultrasound datasets acquired by an automatic sweep over the slightly compressed breast in supine position. This allows for the reconstruction of coronal and sagittal views in addition to the generic transversal view. The selection of a filter for preprocessing is always dependent of the following application. In the context of lesion segmentation an edge enhancement of the lesion border is desired, while smoothing speckle and texture. Features to be preserved

are an irregular margin of the lesion, which is used as indication of malignancy, as well as the architectural distortion and (although rarely seen) microcalcifications inside a lesion. One unique feature found in the coronal view is the star like retraction phenomenon which can appear around malignant lesions. In the context of automated detection, the conservation of this feature is an important requirement for a suitable speckle filter. We have chosen 10 datasets featuring breast lesions differing in size, shape, margin and image quality to compare the filter methods against.

The set of filters we compare here comprise the median filter, the Perona-Malik diffusion filter [4] and the diffusion stick filter [7]. We included a two dimensional (3×3) and three dimensional ($3 \times 3 \times 3$) median filter to evaluate the effect of the neighboring slices on the filter result. The well known Perona-Malik diffusion filter [4] is a nonlinear diffusion filter, where the strength of the diffusion is controlled by an diffusivity function which acts as an 'edge detection' function. In our implementation we chose the edge parameter $\lambda = 2$, the smoothing value $\sigma = 0.1$, a time step of $\Delta t = 1$ and 10 iterations. The idea behind this method is to eliminate small gray value differences which are due to noise and to retain large differences due to signal. In the presence of speckle noise however, gray values differences due to noise can be larger than those due to signal, which can hamper the filter result.

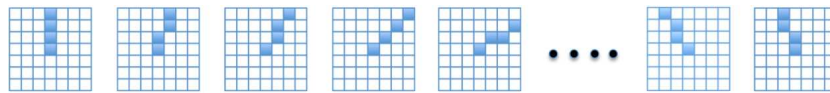


Fig. 1. This image shows 7 examples of the 24 sticks for a 7×7 kernel.

The diffusion stick method [4] is less well known and shall be briefly explained here. It combines the anisotropic diffusion with the works of Czerwinski et al. [9]. A kernel of size $N \times N$ (N must be odd) can be split into $4N - 4$ radial sticks each with the length $(N + 1)/2$. Each stick starts at the kernel center and ends at one of the $4N - 4$ boundary voxels (Fig. 1). Variance and mean along each stick are calculated. The variance acts as edge detection function and controls the degree of diffusion of the mean voxel value. This way, voxel values along lines and edges are preserved while smoothing homogeneous regions. Additionally, a stopping function σ_2 manually set is included to control the overall smoothing (eq. (13) in [7] for details). As we chose a kernel of 7×7 , there will be 24 sticks with the size 4. We also included a three-dimensional diffusion stick method with 218 sticks inside a $7 \times 7 \times 7$ kernel. The stopping function σ_2 is set to 100. The algorithms mentioned above have been implemented in MeVisLab 2.1a and were tested on a 2.8 GHz Quad-Core Intel Xeon Mac (Snow Leopard) with 8 GB RAM.

3 Results

We evaluated each filter on 10 lesions with differing characteristics. Two of these lesions are shown in Fig. 2(a, b). The result of all evaluated filters are shown in

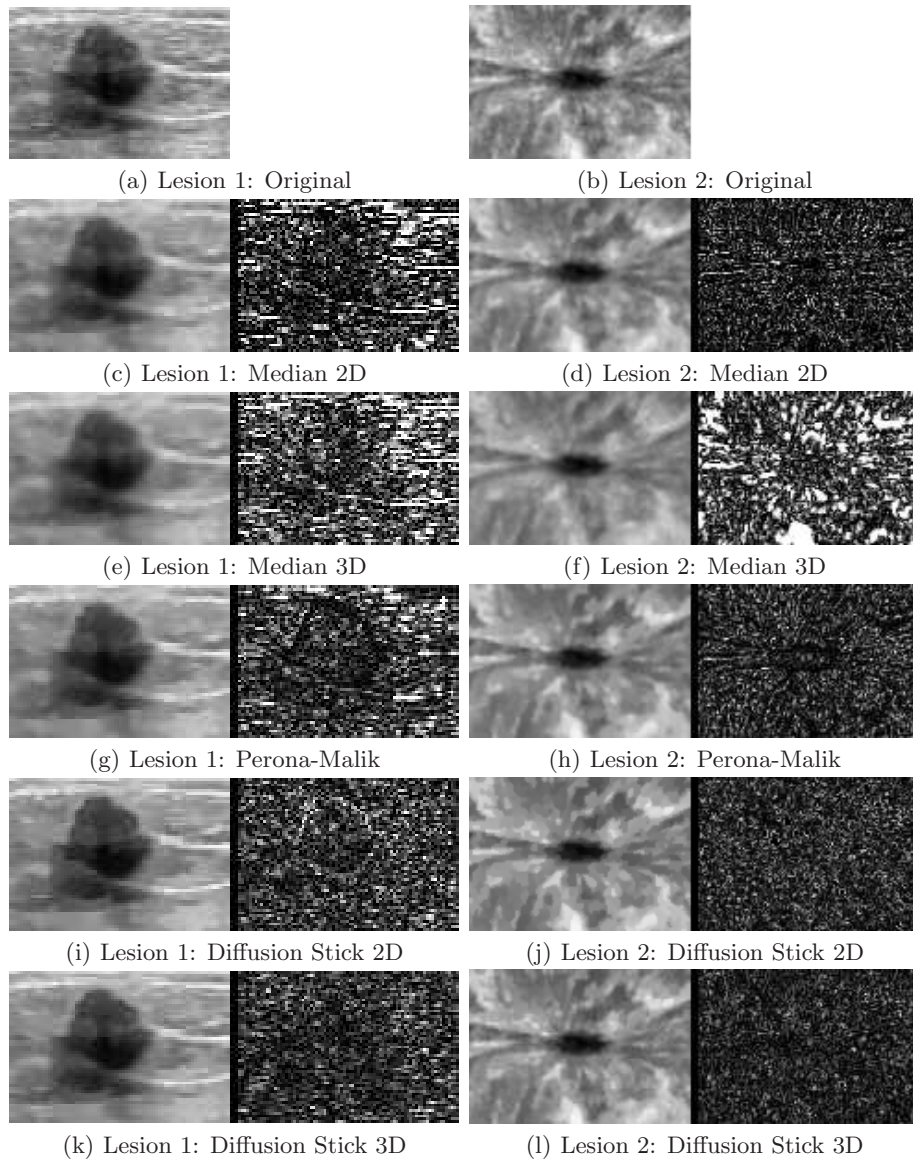


Fig. 2. (a) Lesion 1 features a lobulated shape and possible calcification inside shown in transversal view. (b) Lesion 2 features a retraction the shown in the coronal view. (c-l) The result of each evaluated filter to the respective original lesion images above is shown on the left hand of each subfigure together with the absolute difference to the corresponding original image on the right hand. The difference images visualizes the effect of each filter.

Table 1. Computation time of filters. The whole data block contains 330 slices.

Filter	single slice in [ms]	data block in [s]
Median 2D	3.4	1.1
Median 3D	13.8	4.6
Perona-Malik	82.0	27.1
Diffusion Stick 2D	36.8	12.3
Diffusion Stick 3D	158.2	52.2

the Fig. 2(c-l) below the respective original image. To make the effect of a filter easier visible, each Fig. 2(c-l) shows the difference image of the filter results to the original image on the right hand.

Additionally to the qualitative filter result we compared the computational costs of each filter type in our implementation for a three-dimensional ultrasound dataset of $326 \times 99 \times 330$ voxels. Tab. 1 shows the computation time for a single slice and the whole dataset. The two-dimensional implementation of the median filter needed 1.1 seconds for completion, while the three-dimensional version needed 4.6 seconds. The computation for the Perona-Malik diffusion filter took 27.1 seconds. The two-dimensional diffusion stick algorithm computed the dataset in 12.3 seconds while the three-dimensional implementation needed 52.2 seconds due to the large amount of sticks.

4 Conclusions

The median filter has the lowest computational costs, but its edge preserving capabilities are much weaker compared to the other methods. Moreover, the microcalcifications are almost lost in the filter process as can be seen in Fig. 2(c). Median filtering also appears to thicken or even eliminate linear image features. This is best observed through the respective difference images in the Fig. 2(c, d) and is especially strong in the three-dimensional implementation (Fig.2(e, f)).

With our chosen parameters Perona-Malik diffusion showed better results than the median filter, but (as can be seen in the difference images in Fig. 2(g, h)) there seems to be less filter effect at the position of the lesion borders. This means that the lesion edges do not get smeared while smoothing speckle, but they also do not get enhanced much. Additionally the Perona-Malik diffusion filter appeared to be rather sensitive on the chosen parameters in our experiments and getting pleasing results often needed fixing of the parameters when changing between different datasets of the same quality.

The diffusion stick filter is not as fast as the median filter but as a whole it showed the best results for preserving the features described above and for enhancing the edges of the lesions. Compared to the two-dimensional version, the three-dimensional implementation of the diffusion stick method does not seem to have an improving effect on the filter result, which would compensate

for its large computational time. Edges seemed not to get as much enhanced compared to the two-dimensional implementation which is best seen in Fig. 2(i).

Due to the combination of the good results and its comparable low computational costs we chose the two-dimensional diffusion stick filter as our preprocessing filter of choice.

Future work will concentrate on an evaluation of different speckle reduction techniques with respect to specific image analysis tasks in breast ultrasound, such as segmentation of lesions and anatomical structures, feature extraction and classification, and image registration.

References

1. Burckhardt CB. Speckle in ultrasound B-mode scans. *IEEE Trans Sonics Ultrasonics*. 1978;25(1):1–6.
2. Wagner RF, Smith SW, et al. Statistics of speckle in ultrasound B-scans. *IEEE Trans Sonics Ultrasonics*. 1983;30(3):156–63.
3. Vue Y, Croitoru MM, et al. Nonlinear multiscale wavelet diffusion for speckle suppression and edge enhancement in ultrasound images. *IEEE Trans Med Imag*. 2006;25(3):297–311.
4. Perona P, Malik J. Scale space and edge detection using anisotropic diffusion. *IEEE Trans Pattern Anal Mach Intell*. 1990;12(7):629–39.
5. Yu Y, Acton ST. Speckle reducing anisotropic diffusion. *IEEE Trans Image Processing*. 2002;11:1260–70.
6. Tauber C, Batatia H, Ayache A. A robust speckle reducing anisotropic diffusion. *IEEE Int Conf Image Process*. 2004;1:247–50.
7. Xiao CY, Zhang S, Chen YC. A diffusion stick method for speckle suppression in ultrasonic images. *Pattern Recognit Lett*. 2004;25(16):1867–77.
8. Avianto E, Ito M. Speckle reduction for ultrasonic images using fuzzy morphology. *Trans Inf Syst*. 2001;84(4):502–10.
9. Czerwinski RN, Jones DL, OBrien WD. Detection of lines and boundaries in speckle images- application to medical ultrasound. *IEEE Trans Med Imaging*. 1999;18(2):126–36.

Schätzung der Midsagittalebene zur Bestimmung der Seitenlage maligner Strukturen des Halses

Ivo Rössling¹, Peter Hahn², Lars Dornheim^{1,2}

¹Otto-von-Guericke-Universität Magdeburg

²Dornheim Medical Images

`iroess@isg.cs.uni-magdeburg.de`

Kurzfassung. Für die Bewertung von Tumoren des Halses spielt neben Kriterien wie Größe, Infiltration und anatomischem Bezirk vor allem auch die Seitenlage des Tumors sowie lokaler Metastasen eine wichtige Rolle. Eine automatische Sondierung der Seitenlage kann das Tumor-Staging stark beschleunigen. Ausgehend von Vorarbeiten Dritter präsentieren wir ein angepasstes Verfahren zur Schätzung der Lage und Ausrichtung der Midsagittalebene basierend auf Oberflächennetzen. Die Eigenschaft des Ergebnisses im Rahmen einer TNM-Klassifikation wird anhand einer darauf basierenden automatischen Bestimmung der Seitenlage segmentierter Lymphknoten oder Tumore überprüft.

1 Einleitung

In der chirurgischen Diagnostik und Therapieplanung spielt die Beurteilung anatomischer und pathologischer Strukturen mittels bildgebender Verfahren eine tragende Rolle. Auf der Suche nach Abnormalitäten nutzen Mediziner ihr über Jahre erworbenes Expertenwissen, um in einer vorliegenden Aufnahme einer Region von Interesse Diskrepanzen zum typischen Erscheinungsbild in der jeweiligen Modalität auszumachen. Derartige Abweichungen zeigen sich dabei vor allem in einer veränderten Färbung, Homogenität oder Form. Aber auch atypische Dissymmetrien können auf mögliche Anomalien hindeuten.

Die Median- bzw. Midsagittal-Ebene stellt hierbei eine besondere Referenz dar, vor allem im Rahmen einer verstärkt computerunterstützten Diagnostik. Sie kann zum Beispiel zur Ausrichtung entlang eines Bezugssystems (wie etwa den Talairach-Rahmen) verwendet werden oder als Basis einer Registrierung (gleichartiger Strukturen oder gegen einen Atlas) dienen. Sie eignet sich auch Ausgangspunkt einer Dissymmetrie-Analyse, wie sie z.B. [1] durchgeführt wird, um automatisch mögliche Tumor-Areale im Bereich des Kopfes zu lokalisieren. Im Kontext segmentierter anatomischer bzw. pathologischer Strukturen genügt bereits eine gute Schätzung zur Lage dieser Ebene, um in vielen Fällen die Seitigkeiten einzelner Strukturen vollautomatisch bestimmen zu können. Dies ist insbesondere im Rahmen des Tumor-Staging (TNM-Klassifikation) von besonderer Bedeutung, da die jeweilige Seitenlage befallener Lymphknoten im Vergleich zu der des Tumor (unilateral, bilateral, ipsi- oder kontralateral) entscheidenden Einfluss auf die Behandlungsmöglichkeiten hat. Dieser Anwendungsfall soll im Fokus der folgenden Untersuchungen stehen.

2 Verwandte Arbeiten

In der Literatur der allgemeinen und der medizinischen Bildverarbeitung wurde die Frage der Bestimmung von Symmetrien wurde schon häufiger aufgegriffen. Im medizinischen Kontext stand dabei hauptsächlich das Gehirn im Zentrum des Interesses. Eine gute Übersicht zu diesem Thema bietet [2]. Für die Bestimmung einer derartigen Symmetrieebene können allgemein zwei Strategien unterschieden werden, die sich aus ihrer gleichzeitigen anatomischen und symmetrischen Natur ergeben. Zum einen kann eine Suche nach Landmarken oder anderen morphologische Merkmalen mit anschließendem Fitting durchgeführt werden. Mittels Hough-Transformationsprinzip wird so z.B. in [3] der Interhemisphärenspalt detektiert. Zum anderen kann das Maß die Ermittlung einer Ebene sein, die ein geeignetes Symmetriekriterium (weitestgehend) maximiert. Übliche Ansätze sind nicht-rigide Registrierung eines (3D-)Datensatzes gegen sein Spiegelbild [2] oder Matching durch Maximierung der Kreuz-Korrelation unter Rotation und Translation [4]. Auch Histogramme fanden als Kriterium Anwendung [1]. Als Daten dienten jeweils Grauwert- oder Kantenbilder.

Alternative Methoden reichen von statistischen Maßen [5] über algebraische Ansätze [6] bis hin zu generalisierten Symmetrietransformationen [7] und Symmetrie-Deskriptoren [8]. Die Idee, Symmetrien im (2D-)Bildraum über Symmetrien im Gradienten-Histogramm zu identifizieren, wurde erstmalig in [9] präsentiert. Entlang der Symmetrieachse müssten im optimalen Falle die beiden Hälften des Winkelhistogramms einander entsprechen. Ein Least-Squares-Fitting des Histogramms mit seinem schrittweise rotierten Spiegelbild sollte demnach eine gute Näherung der Orientierung der Spiegelachse liefern. Diese Idee lässt sich nicht nur auf die dritte Dimension erweitern, sondern lässt sich auch auf Meshes übertragen [10], doch wird für keinen dieser Anwendungsfälle die Qualität oder Verwertbarkeit des Ergebnisses im Anschluss näher betrachtet.

3 Material und Methoden

Für unsere Untersuchungen nutzten wir Segmentierungen des Halses in Form von Oberflächen-Dreiecksnetzen. Voxelbasierte Segmentierungen könnten in diesem Zusammenhang ohne komplexitätsmäßigen Mehraufwand in äquivalente Dreiecksnetze überführt werden. Ziel soll es sein, anhand der Segmentierung der Knochen eine Schätzung der Midsagittalebene durchzuführen. Für diese soll im Anschluss die korrekte automatische Beurteilung der Seitenlage der restlichen segmentierten Strukturen überprüft werden.

Zur Berechnung der Midsagittalebene orientieren wir uns an dem in [10] beschriebenen Verfahren. Zunächst werden mittels einer Hauptachsenanalyse der Schwerpunkt sowie die Hauptachsen der segmentierten Knochen bestimmt. Dieser Schritt ist dadurch motiviert, dass im (theoretischen) Falle perfekter Symmetrie gilt: Jede Symmetrieebene eines Körpers ist orthogonal zu einer seiner Hauptachsen – und jede Symmetrieachse eines Körpers ist zugleich eine Hauptachse desselbigen [11]. Zunächst ist jedoch unklar, welche die in diesem Sinne qualifizierte Hauptachse ist.

Um die Orientierung der Symmetrieebene zu ermitteln werden in [10] als nächstes die Flächennormalen des Oberflächennetzes auf der Einheitskugel abgetragen, welche aufgrund des nicht notwendigerweise homogen vernetzten Meshes jeweils mit dem korrespondierenden Flächeninhalt gewichtet sind. Es entsteht das sogenannte „Extended Gaussian Image“ (EGI, Abb. 1), welches anschließend zu einem 3D-Orientierungshistogramm diskretisiert wird. Der Mangel an einer geeigneten Symmetriegruppe beliebig großer Ordnung bezüglich Spiegelung auf der Einheitskugel im \mathbb{R}^3 , verwehrt jedoch die Möglichkeit, dieses 3D-Histogramm direkt gegen ein rotiertes Spiegelbild seiner selbst zu matchen, wie es im \mathbb{R}^2 möglich ist. Stattdessen muss für jede Rotation ein neues Histogramm aus dem EGI berechnet werden. Um den Berechnungsaufwand gering zu halten, werden in [10] an dieser Stelle nun die drei Hauptachsen dazu herangezogen, die globale Suche der Orientierung der Symmetrieebene auf eine lokale Suche in der Nähe der Hauptachsen zu reduzieren. Hierzu werden die jeweils fünf bis sechs benachbarten Zellen des Histogramms untersucht.

Für unsere Untersuchungen haben wir das Verfahren aus [10] jedoch in zwei grundsätzlichen Punkten angepasst. Zum einen ziehen wir statt der Flächennormalen vielmehr die Punktnormalen des Oberflächennetzes zur Bildung des Histogramms heran. Als individuelle Wichtung wird die Summe der Flächeninhalte aller jeweils angrenzender Flächen benutzt. Die Präferenz der Punktnormalen gegenüber Flächennormalen rührt daher, dass bei bestimmten Verfahren der Meshherzeugung (z.B. Marching Cubes über Binärdaten) die Flächennormalen systembedingt stark eingeschränkt sind in ihrer Ausrichtung. Punktnormalen dagegen zeigen ein weitaus homogenere Verteilung.

Darüber hinaus haben wir uns gegen die Verwendung von 3D-Histogrammen entschieden. Stattdessen führen wir zunächst basierend auf den Hauptachsen eine Basistransformation des Oberflächennetzes durch. Für die sich ergebende Repräsentation werden anschließend drei 2D-Orientierungshistogramme (bezüglich der XY-, XZ- und YZ-Ebene) generiert. Ein Least-Squares-Matching eines jeden solchen Histogramms mit seinem schrittweise rotierten Spiegelbild (Abb. 1) liefert die Ausrichtung der gesuchten Spiegelebene.

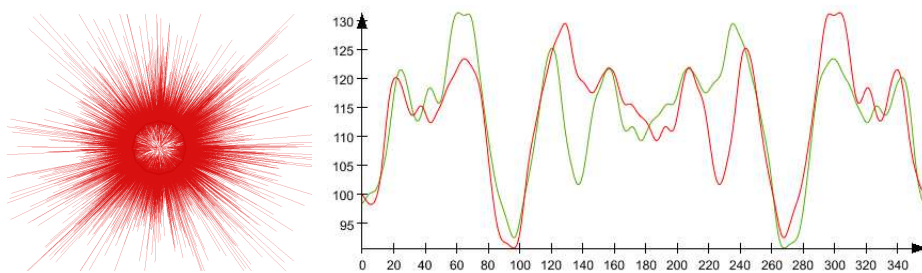


Abb. 1. *Links:* 3D Extended Gaussian Image für einen der untersuchten Datensätze. *Rechts:* Least-Squares-Matching eines 2D-Winkelhistogramms mit seinem gedrehten Spiegelbild für denselben Datensatz für das gefundene Optimum bei 5° .

4 Ergebnisse

Zur Evaluierung des Verfahrens wurde eine Fallstudie basierend auf 10 Datensätzen segmentierter Strukturen des Halses durchgeführt. Neben Knochen, Tumor und Lymphknoten waren noch weitere Strukturen wie Blutgefäße, Muskeln, Trachea oder Ring-/Schilddrüse enthalten, wurden mangels Relevanz für den Anwendungsfall allerdings nicht näher betrachtet. Die Segmentierungen wurden durch den klinischen Kooperationspartner im Vorfeld erstellt und lagen in Form von Oberflächendreiecksnetzen vor.

Für jeden der Datensätze konnte in weniger als drei Sekunden die gewünschte Midsagittalebene ermittelt sowie Seitigkeit des Tumors und der Lymphknoten bestimmt werden. Insgesamt wurden 10 Tumore und 134 (nicht zwingend befallene) Lymphknoten auf ihre Seitigkeit getestet. In drei Fällen wurde ein Lymphknoten als der durch den Arzt festgelegten Seite gegenüberliegend klassifiziert. In diesen Fällen wurde die Lage zur Midsagittalebene allerdings durch den Experten auch als grenzwertig nahe beurteilt. Viermal wurde für den Tumor korrekt eine Überschreitung der Mittellinie festgestellt, ein weiteres Mal entschied der Arzt dagegen im Vorfeld auf einen noch einseitigen Tumor. Eine Festlegung auf eine Primärseite wurde durch unser Programm generell nicht weiter vorgenommen. In zwei Fällen wurde ein Lymphknoten entgegen ärztlicher Bewertung als mittig klassifiziert. Auf die TNM-Klassifikation hatte dies jedoch keinen Einfluss, da mittig gelegene befallene Lymphknoten noch als ipsilateral angesehen werden.

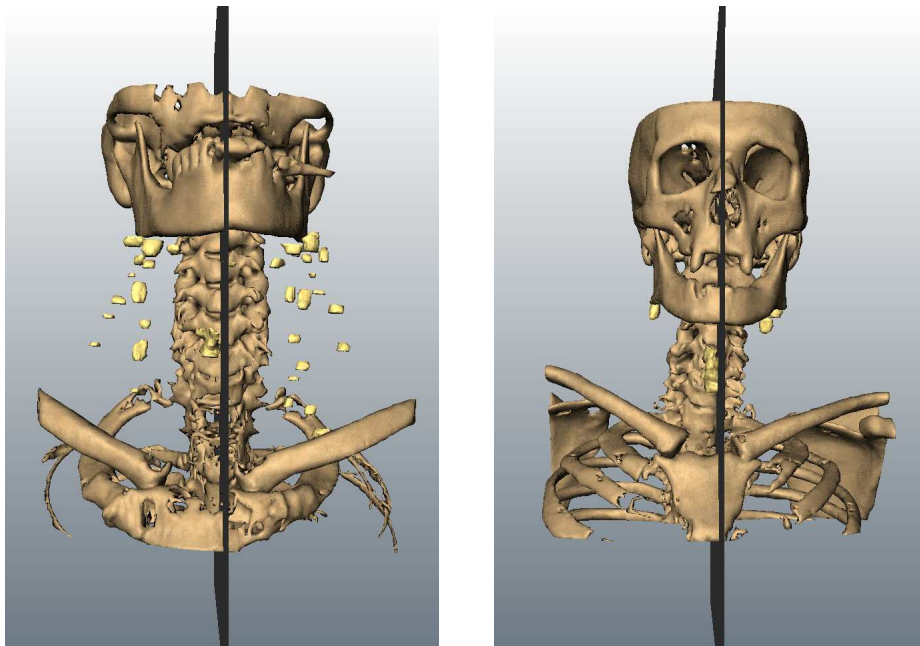


Abb. 2. Ergebnis der Midsagittalebene-Schätzung für zwei untersuchte Datensätze.

5 Diskussion

Es wurde ein Verfahren vorgestellt, dass basierend auf einer als Oberflächen-netz gegebenen Segmentierung des Kopfes und der Halsknochen eine Schätzung der Midsagittalebene durchführt. Primärintention war die anschließende automatische Seitenlagenbestimmung von Halstumoren und Lymphknoten, wie sie im Rahmen des Tumor-Stagings erhoben wird. In empirischen Tests an 10 klinischen Datensätzen wurde festgestellt, dass die Qualität des Ergebnisses in diesen Fällen stets hinreichend war, um eine korrekte Lateralität zu ermitteln bzw. auf Grenzfälle in Form einer Mittellinienüberschreitung hinzuweisen. Keiner der Einzelfälle mit abweichender Expertenmeinung führte letztlich zu einer veränderten Gesamteinschätzung, welche die Lateralität des Befalls über alle malignen Lymphknoten für die TNM-Klassifikation subsumiert. Der mögliche Fehler durch Einschränkung der Orientierungshistogramme auf 2D scheint vernachlässigbar. Damit zeigt sich das Verfahren (auch mit Hinblick auf die benötigte Laufzeit) vom Grundsatz her für die Anwendung geeignet.

Zudem bleibt der effektive Fehler des beschriebenen Verfahrens noch detailliert zu untersuchen und zu quantifizieren. Eine Erhebung der mittleren Abweichung von einer jeweils durch einen medizinischen Experten angegebenen Ebene in Position und Orientierung ist derzeit in Bearbeitung. Insbesondere ist offen und bleibt somit noch zu untersuchen, wie robust oder sensibel sich das Verfahren letztlich im Falle hinreichend schlechter Datenqualität erweist.

Literaturverzeichnis

1. Mancas M, Gosselin B, Macq B. Fast and automatic tumoral area localisation using symmetry. In: Proc IEEE ICASSP. vol. 2; 2005. p. 725–8.
2. Prima S, Ourselin S, Ayache N. Computation of the mid-sagittal plane in 3-D brain images. *IEEE Trans Med Imaging*. 2002;21(2):122–38.
3. Brummer ME. Hough transform detection of the longitudinal fissure in tomographic head images. *IEEE Trans Med Imaging*. 1991;10:74–81.
4. Liu Y, Collins R, Rothfus WE. Robust midsagittal plane extraction from normal and pathological 3-D neuroradiology images. *IEEE Trans Med Imaging*. 2001;20(1):175–92.
5. Thirion JP, et al. Statistical analysis of normal and abnormal dissymmetry in volumetric medical images. In: Proc IEEE WBIA; 1998. p. 74.
6. Keller Y, Shkolnisky Y. An algebraic approach to symmetry detection. In: Proc IEEE ICPR. vol. 3; 2004. p. 186–9.
7. Reissfeld D, Wolfson H, Yeshurun Y. Context free attentional operators: The generalized symmetry transform. *Int J Computer Vis*. 1995;14:119–30.
8. Kazhdan M, Chazelle B, Dobkin D, et al. A Reflective symmetry descriptor for 3d models. *Algorithmica*. 2003;38(1):201–25.
9. Sun C. Symmetry detection using gradient information. *Pattern Recognit Lett*. 1995;16(9):987–96.
10. Sun C, Sherrah J. 3D Symmetry detection using the extended gaussian image. *IEEE Trans Pattern Anal Mach Intell*. 1997;19(2):164–8.
11. Minovic P, Ishikawa S, Kato K. Three-Dimensional symmetry identification part I: Theory. *Memoirs of the Kyushu Institute of Techn Eng*. 1992;21:1–16.

GPU-basiertes Volumenrendering von multimodalen medizinischen Bilddaten in Echtzeit

Ingrid Scholl¹, Nicole Schubert¹, Pascal Ziener¹, Uwe Pietrzyk²

¹Fachbereich für Elektrotechnik und Informationstechnik, FH Aachen

²Institut für Neurowissenschaften und Medizin, Forschungszentrum Jülich
scholl@fh-aachen.de

Kurzfassung. Die vorliegende Arbeit zeichnet sich dadurch aus, dass registrierte unsegmentierte Volumina aus multimodalen Bilddatensätzen (z.B. MRT, PET) gleichzeitig in einer 3D-Rekonstruktion visualisiert werden und in Echtzeit manipuliert werden können. Ermöglicht wird die Echtzeitfähigkeit durch die Programmierung der Algorithmen zur direkten Volumenvisualisierung auf der Grafikkarte mittels der neuen CUDA-Technologie. Die Zuordnung der Farbeigenschaften wird über 1D-Transferfunktionen für jedes Volumen getrennt gesteuert. So können durch die interaktive Veränderung der 1D-Transferfunktion Detailinformationen aus den zwei Bilddatensätzen getrennt kontrolliert werden und die Vorteile der verschiedenen Bildmodalitäten in einer Visualisierung genutzt werden. Mittels dieses interaktiven *Frameworks* können neue Erkenntnisse insbesondere über neurodegenerativen Erkrankungen gewonnen werden.

1 Einleitung

Zur Diagnostik neurodegenerativer Erkrankungen des menschlichen Gehirns sowie weiterer Tumorerkrankungen werden mehrere bildgebende Untersuchungen am gleichen Patienten durchgeführt, um die Vorteile der verschiedenen bildgebenden Verfahren zu nutzen. Die gewonnenen multimodalen Bilddatensätze aus MRT, PET und CT werden in der Regel einzeln betrachtet und vergleichend bewertet. Eine 3D Rekonstruktion der Bilddatensätze liefert dem Mediziner zusätzliche Informationen und kann durch die Verfahren der direkten Volumenvisualisierung gewonnen werden. Wenn eine höhere Qualität bei der 3D Visualisierung gewünscht wird, können einfachere Verfahren der Volumenvisualisierung wie 2D- und 3D-Texture-Slicing nicht verwendet werden. In diesem Fall wird das Verfahren der direkten Volumenvisualisierung mittels Ray Casting für uniforme Gitter verwendet. Hierbei wird die Licht-Transferfunktion entlang von Sehstrahlen durch das Volumen bestimmt. Die abgetasteten Werte entlang des Sehstrahls werden trilinear aus den umgebenden Voxel-Werten interpoliert und aufaddiert.

Um Echtzeitfähigkeit zu erreichen, werden die Algorithmen zur Volumenvisualisierung auf die Grafikkarte ausgelagert [1, 2]. Dazu stehen C-ähnliche Programmiersprachen, wie CUDA von Nvidia zur Verfügung, um beliebige parallele

Algorithmen mit Hilfe der grafischen Prozessoreinheit (GPU/GPGPU) anstatt der CPU zu berechnen.

Schwierig wird es, wenn mehrere Volumendaten gleichzeitig in einem Volumen abgebildet werden sollen, die zunächst durch eine Vorverarbeitung registriert werden müssen. Die Verfahren zur multimodalen Visualisierung durch das Ray Casting unterscheiden sich hierbei in der Literatur, ob nur eine Transferfunktion für alle Volumina verwendet wird, oder ob eine Transferfunktion pro Volumen eingesetzt wird. Durch die Verwendung von 1D-Transferfunktionen können Farbe und Transparenz durch die Kombination der abgetasteten Ray Casting Strahlenwerte gesteuert werden. Verschiedene Techniken der Farb- und Transparenzzuordnung werden in [3] vorgestellt.

2 Material und Methoden

Gegeben sind *rigide* registrierte MRT und CT sowie MRT und PET Bilddatensätze von mehreren Untersuchungen des menschlichen Gehirns und Körpers der Größe 512^3 oder 256^3 .

Die Abbildung 1 zeigt den von uns implementierten Ansatz: Nachdem die Bilddaten aufeinander registriert wurden, werden für jedes Volumen 1D-Transferfunktionen bestimmt. Die Volumendaten sowie deren Transferfunktionen werden auf die GPU transferiert. Dort wird pro Volumen ein direktes Volumenrendering mittels Ray Casting durch das Front-to-Back-Verfahren in Echtzeit berechnet, wobei C_{src} der abgetastete Farbwert auf dem Sehstrahl entspricht, α_{src} der dazugehörige Opazitätswert nach der 1D-Transferfunktion und C_{dst} sowie α_{dst} die entsprechenden akkumulierten Werte entlang des Sehstrahls [4]:

$$C_{dst} \leftarrow C_{dst} + (1 - \alpha_{dst})\alpha_{src}C_{src} \quad (1)$$

$$\alpha_{dst} \leftarrow \alpha_{dst} + (1 - \alpha_{dst})\alpha_{src} \quad (2)$$

Das Ray Casting wurde mit dem Verfahren der Early Ray Termination optimiert, so dass die Abtastwerte nicht weiter in die Berechnung einfließen, wenn der akkumulierte α_{dst} -Wert einen vorgegebenen Schwellwert t erreicht. Den Ergebnissen pro Pixel werden über die volumenspezifische Transferfunktion eine

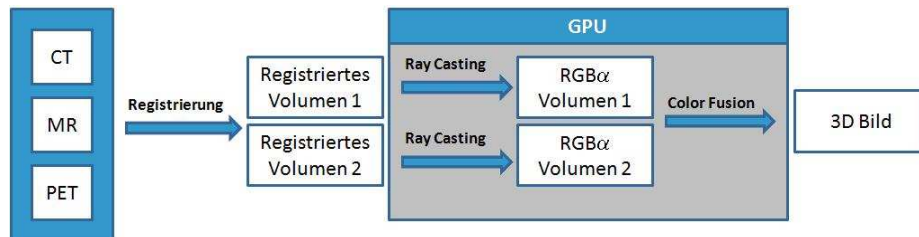


Abb. 1. Workflow der Implementierung.

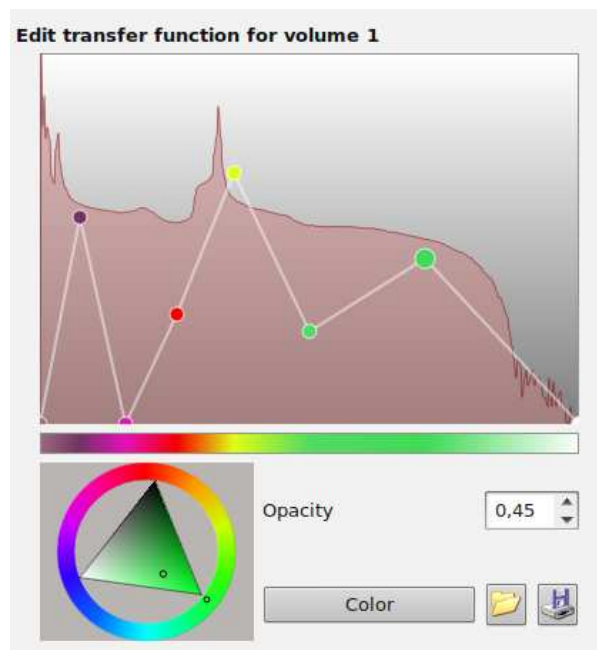
Farbe und eine Transparenz zugeordnet, so dass man pro Volumen und Pixel jeweils einen $RGB\alpha$ -Wert erhält. Durch das Verfahren der Color Fusion wird ein kombinierter Farbwert aus den einzelnen Farbwerten der multimodalen Volumen pro Pixel zugeordnet, der letztendlich als gewichtete Summe aus diesen für die Farbe und die Transparenz berechnet und dargestellt wird [3]:

$$C_{dst} = \beta C_{dst,V1} + (1 - \beta) C_{dst,V2} \quad (3)$$

$$\alpha_{dst} = \beta \alpha_{dst,V1} + (1 - \beta) \alpha_{dst,V2} \quad (4)$$

Die Transferfunktionen können interaktiv mit Hilfe der GUI aus Abb. 2 editiert werden. Die Änderungen werden in Echtzeit auf die Datensätze übertragen und in der 3D Visualisierung direkt sichtbar.

Abb. 2. GUI zur Definition der 1D-Transferfunktion pro Volumen. Im Hintergrund ist das Histogramm der Ray Casting Strahlen zu einem Volumen abgebildet. Die Farben wie auch die Opazität können vom Benutzer interaktiv gesetzt werden und bewirken eine direkte Auswirkung auf das multimodale Visualisierungsergebnis.



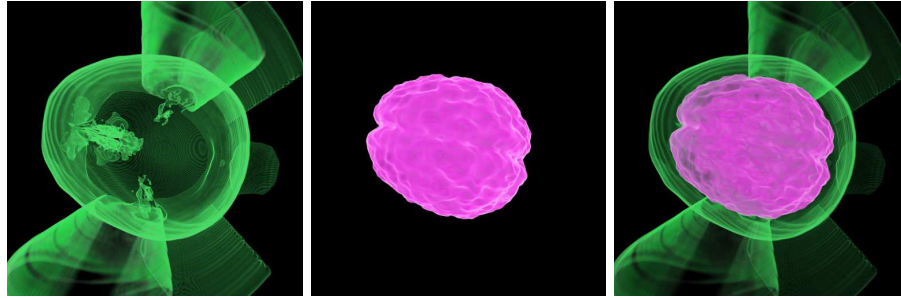
3 Ergebnisse

Der Algorithmus wurde auf einem System mit Intel Core i7 920 Prozessor mit 2.67 GHz, einer Nvidia Quadro FX 5800 Grafikkarte mit 4 GB Grafikspeicher und 12 GB RAM implementiert. Das Programm ist unter Ubuntu 9.04 64 bit und Windows XP 64 bit lauffähig.

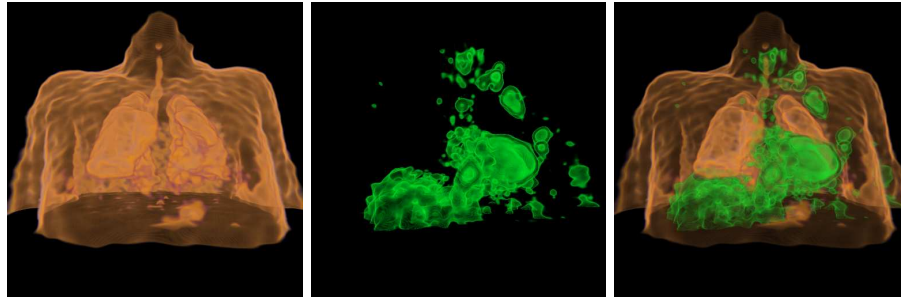
Die Visualisierungsergebnisse sind in Abb. 3 zu sehen. In allen Darstellungen sind die Datensätze einzeln und zusammen veranschaulicht mit unterschiedlichen Transferfunktionen mit $\beta = 0.5$ nach Gl. 4. Es ist sehr gut zu sehen, dass

der Informationsgehalt aus den Datensätzen in der gemeinsamen Visualisierung zunimmt (Abb. 3b,c).

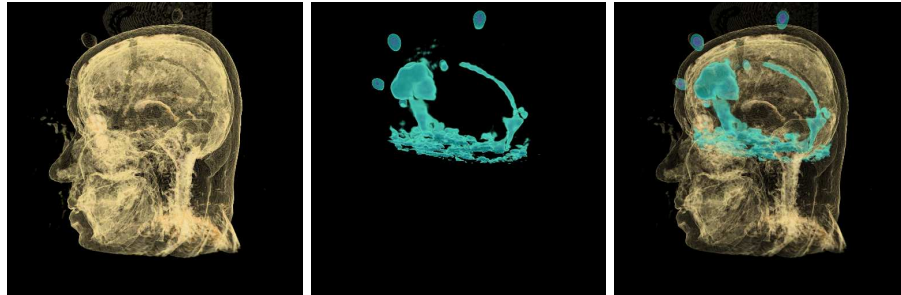
Der Algorithmus wurde daraufhin untersucht, wie sich die Bildwiederholrate pro Sekunde (fps) in Abhängigkeit der Auflösung des gerenderten Ausgabebildes und der Anzahl der maximalen Abtastwerte für alle Ray Casting Strahlen verändert. In Tab. 1 ist das Ergebnis dargestellt. Es ist deutlich, dass mit steigender Anzahl der Abtastwerte die Bildwiederholrate pro Sekunde abnimmt.



(a) Visualisierung von CT und PET Datensätzen des menschlichen Kopfes



(b) Visualisierung von 2 PET Datensätzen des menschlichen Thorax



(c) Visualisierung von MRT und PET Datensätzen des menschlichen Kopfes

Abb. 3. Visualisierung von multimodalen Datensätzen der Größe 512×512 . Links: erster Datensatz; Mitte: zweiter Datensatz; rechts: multimodaler Datensatz.

Tabelle 1. Messwerte: Framerate pro Sekunde für die maximale Anzahl von Abtastwerten (berechnet aus Schrittweite der Abtastwerte des Sehstrahls und Volumengröße).

Schrittweite	Bildgrösse	Max. Abtastwerte	fps
0.01	512×512	52.428.800	82
0.005	512×512	104.857.600	42
0.01	1024×1024	209.715.200	26
0.005	1024×1024	419.430.400	16
0.001	512×512	524.288.000	12
0.001	1024×1024	2.097.152.000	5

4 Diskussion

Der vorgestellte Algorithmus bietet eine neue Applikation für die multimodale Visualisierung medizinischer Bilddaten in Echtzeit. Durch die interaktive Steuerung der Transferfunktionen der Volumen können die Eigenschaften beider Volumina visualisiert werden. Ohne Echtzeitfähigkeit ist das richtige Setzen der Transferfunktionen ein mühsamer Prozeß, der hier durch die neue Technik der GPU Programmierung umgangen werden konnte. Um die Methode weiter zu optimieren, ist die Nutzung von zusätzlichen Grafikkarten vorgesehen, um auch große Volumen darzustellen. Dazu wird auch die Nutzung von Komprimierungstechniken, wie z.Bsp. Wavelet Komprimierung, notwendig, um so den Speicherbedarf der Volumen zu reduzieren. Desweiteren werden andere Möglichkeiten implementiert, die die beiden Volumen fusionieren, wie Depth Peeling oder Opazität gewichtetes Mischen der Volumen [5].

Danksagung. Wir danken dem Bundesministerium für Bildung und Forschung für die Unterstützung dieser Arbeit.

Literaturverzeichnis

1. Stegmaier S, Strengert M, Klein T, et al. A simple and flexible volume rendering framework for graphics-hardware-based raycasting. Proc Vol Graph. 2005; p. 187–95.
2. Weiskopf D. GPU-based Interactive Visualization Techniques. Berlin: Springer; 2006.
3. Beyer J, Hadwiger M, Wolfsberger S, et al. High-quality multimodal volume rendering for preoperative planning of neurosurgical interventions. Proc IEEE Vis. 2007; p. 1696–703.
4. Engel K, Hadwiger M, Kniss J, et al. Real-Time Vol Graph. Peters; 2006.
5. Brecheisen R, i Bartroli AV, Platel B, et al. Flexible GPU-based multi-volume ray-casting. Proc VMV. 2008; p. 303–12.

Automatic Liver Segmentation in Contrast-enhanced MRI

René Siewert¹, Dirk Schnapauff², Timm Denecke², Thomas Tolxdorff¹,
Dagmar Krefting¹

¹Institut für Medizinische Informatik, Charité–Universitätsmedizin Berlin

²Institut für Radiologie, Charité–Universitätsmedizin Berlin

`rene.siewert@charite.de`

Abstract. A fully automated method for liver segmentation in contrast enhanced abdominal MRI scans is presented. Liver shape and volume are obtained utilizing a context based approach. Compared to manual segmentation, the obtained liver volumes differ less than 10 %. The mean sensitivity of 0.92 is comparable to other published liver segmentation methods.

1 Introduction

Image based analysis of a patient’s or a potential transplant donator’s liver is vital for diagnosis, therapy and treatment control; and liver segmentation is an active research area. Anyhow, due to the high number of liver shape variants, and the wide range of morphological and structural changes due to liver-affecting disorders and diseases, robust liver segmentation is still a challenging task. Most algorithms are developed for CT-scans as the current goldstandard for liver imaging[1, 2]. In order to reduce the patient’s X-ray exposure, efforts are notable to at least partially replace CT-scans by MRI. These efforts are supported in particular by new MRI contrast agents. The specific contrast agent Gd-EOB (Primovist[®] in Europe or Eovist[®] in the U.S.) allows for MRI based lesion detection and characterization[3] and is a promising alternative for presurgical abdominal scans. As the exact measurement of the liver volume is required for therapy planning, these scans are currently segmented manually using the software provided by the manufacturer of the scanner, as existing liver segmentation algorithms developed for CT are not applicable for MRI. This is due to the fact, that intensity ranges and distribution in MRI differ strongly from CTs. Furthermore the image contrasts change with the use of contrast agents, so even MRI specific liver segmentation algorithms could not be applied to our data [4, 5, 6]. Therefore, we have developed a fully automated liver segmentation method for Primovist[®] enhanced abdominal MRI. A context based approach has been chosen for the segmentation algorithm. The basic idea of context based segmentation is the classification of image areas by subsequently applying a set of predefined rules to the image. These rules are strongly related to the human decisions taken during the respective pattern or object recognition task, and try

to comprise the experts knowledge. Each rule creates a dependency between different classes of the image. These dependencies form the image context. After all rules are applied, the image context should represent a valid segmentation. The main task within this approach is the formulation of a robust set of rules. In practice, a rule compares different features between image objects. Typical features are intensity or shape/size related, as well as comparative features between objects, such as relative location, distance or intensity difference. In order to get a robust initial segmentation, the basic rules are often formulated in the way to avoid false positives, resulting in an incomplete segmentation. Following the assumption, that possible „false negatives“ (FN) are in the vicinity of the found segmentation and show certain similarities to the neighboring „true positives“ (TP), morphological operations have proved to be a good choice to complete the segmentation.

2 Methods

In total, ten abdominal contrast enhanced MRI scans (*GE Signa Excite*, 3d-T1w fat saturated (LAVA) sequences, single breath-hold acquisition, 2 mm slice thickness) of the liver in the hepatobiliary phase (20 minutes after administration), have been taken within the clinical routine. Primovist is specifically uptaken by liver parenchyma (hepatocytes), thus brightening healthy liver tissue in T1-weighted MRI (parenchymal enhancement). Hepatic vessels and lesions with no or minimal hepatocytic function (cysts, metastases, the majority of hepatocellular carcinomas) will remain unenhanced and appear as dark objects within the liver region. Anyhow, the intensity of healthy liver tissue show gradients, appearing darker in the center of the liver, additionally to an overall gradient in B-field direction, often found in MRI. The grey-value range of the liver parenchyma overlaps with other organs, like kidneys, structures of the abdominal wall or sometimes the stomach even in the same slice (Fig. 1).

Due to the high variety of liver shapes, mainly intensity based features are employed, especially strong edges found at the liver boundary; complemented by spatial and local shape information. As the overall intensity gradient in B-Field direction is significant, many features are extracted for each slice separately,

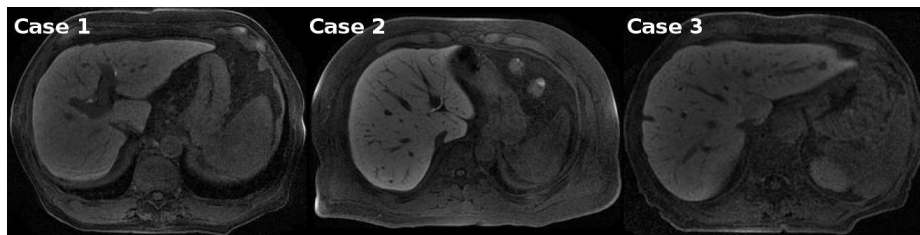


Fig. 1. Contrast enhanced MRT of the liver, different liver shapes.

combining the resulting classes to 3D objects occasionally. The algorithm scheme can be divided into seven steps:

1. Generous intensity based edge detection
2. Knowledge based elimination of false positive liver edges
3. Detection of strong edge and adjacent bright object pairs (liver seeds)
4. Initial liver segmentation
5. Morphological region growing of liver seeds
6. 3D-Reconstruction of found liver
7. Intensity based smoothing of the liver shape

For each slice ($S_{i,o}$), a median filtered ($S_{i,m}$) and a subsequently Sobel filtered derivative ($S_{i,s}$) are generated. Both filter methods use a 3×3 rectangular filter kernel. Edge detection (Step 1) is performed by histogram based thresholding of $S_{i,s}$. The threshold T_1 is found by $T_1 = \text{mean}(P_{95}) + 2/3\sigma(P_{95})$, where P_{95} is the 95 percentile of $S_{i,s}$ and σ denotes the standard deviation. In the subsequent Step 2, false positives of the detected boundary objects are eliminated by evaluation of the position. The pixels belonging to the liver are assumed to be located mainly in the upper-left part of the slices. Therefore objects with the center of mass being found in the lower right quadrant of the slice are excluded. Then bright objects, defined by objects being above threshold $T_2 = \text{mean}(S_{i,m}) + 2/3\sigma(S_{i,m})$, that are connected to the found liver border, are identified as liver parts (Step 3). If no bright objects are found along a particular border, the border itself is excluded from the liver border class. The initial (incomplete) liver segmentation is obtained by taking the largest connected component found in 3D reconstruction of the found 2D liver objects (Step 4). The missing liver parenchyma (Step 5) is now found by again slice based region growing processes of the liver parts. The evaluation of the growing criteria is performed on a further filtered derivative of the MRI scan $S_{i,nb} = 1/2(S_{i,o} + S_{i,m}) - S_{i,s}$. Within this combined image, the overall intensity is slightly smoothed, but strong edges appear as dark borders, and prohibit leakage of the liver parts across the border during the region growing process. The alternative stop criteria for the region growing are given by:

- The candidate pixel's intensity is lower than $\text{mean}(S_{i,nb})$
- The candidate pixel's intensity is lower than the intensity of the respective seed pixel (5% tolerance is considered),
- More than $1/3$ of the pixels within the range 4 Moore neighborhood of the candidate pixel are already belonging to the liver object.

The resulting liver volume is again reconstructed from the found 2D-liver objects (Step 6), and the shape is smoothed by morphological opening with a diamond shaped structuring element of range 3 (Step 7).

3 Results

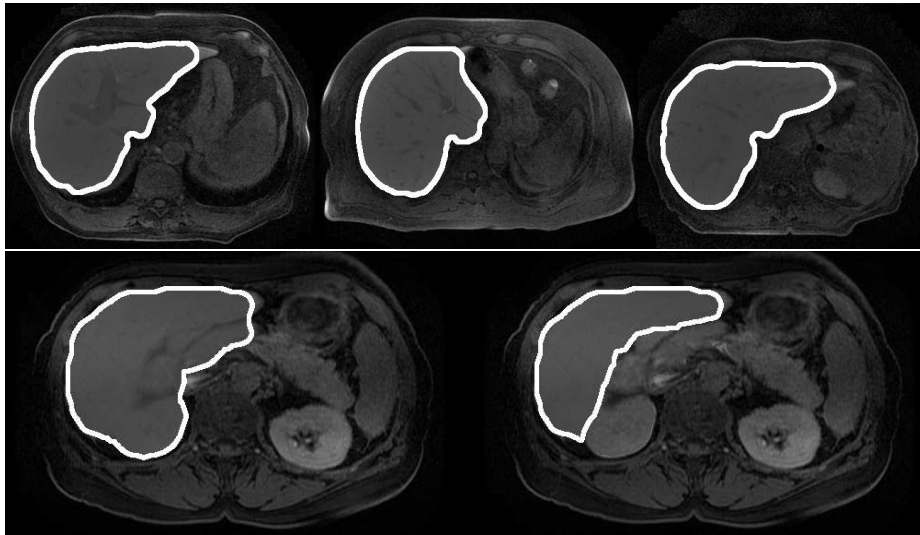
The proposed method is applied to ten clinical datasets of healthy patients as well as patients suffering from liver cancer. The results obtained are compared to manual segmentation provided by the attending radiologist. As

Table 1. Segmentation results. aVol = automatic Volume; mVol = manual Volume.

Case	1	2	3	4	5	6	7	8	9	10	mean	σ
mVol [ml]	1671	1750	1380	1904	1626	1026	1341	1427	1377	2666		
aVol [ml]	1601	1647	1345	1779	1718	1085	1212	1417	1405	2665		
Δ Vol [%]	-4,18	-5,90	-2,53	-6,59	+5,64	+5,73	-9,61	-0,68	+2,04	-0,06	$\pm 4,2$	3,0
Sens	0,931	0,914	0,924	0,900	0,949	0,930	0,854	0,940	0,953	0,934	0,92	0,041

quality measures, the deviation Δ Vol of the automatically obtained liver volumes (aVol) to the manual result (mVol), as well as the pixel-wise sensitivity $\text{Sens} = TP/(TP + FN)$, are chosen. The results are given in Tab. 1; Fig. 2 show exemplary segmentation results.

The measured liver volumes span the wide range of $\text{Vol}_{\min} = 1026$ ml to $\text{Vol}_{\max} = 2666$ ml. All deviations in liver volumetry are within an acceptable tolerance of 10 %. The mean Volume error is $\text{mean}(\Delta\text{Vol}) = \pm 4.2\%$ with $\sigma(\Delta\text{Vol}) = 3,0\%$. The mean sensitivity is found to be $\text{mean}(\text{Sens}) = 0.92$ with $\sigma(\text{Sens}) = 0,041$. These results are comparable to those found for published segmentation methods, e.g. $\text{Sens} = 0.92$ [7], $\text{Sens} = 0.96$ and $\Delta\text{Vol} = 3\%$ [8] for CT segmentation. For the two reported fully automatic MRI based liver segmentation methods, no comparable measures are available. Massoptier et al. give a mean sensitivity $\text{Sens} = 0.95$ for ten datasets encompassing CT and MRI scans [6], while Logeswaran et al. do not give any quantitative measures at all [4]. The obtained results are promising, however, a critical point within the

**Fig. 2.** Segmentation results for different liver shapes (top) inaccurate result including kidney (bottom left), and corresponding manual segmentation (bottom right).

segmentation turned out to be the correct exclusion of the right kidney. In fact the right kidney may be closely attached to the liver, and show similar intensity values. An example is given in the bottom row of Fig. 2).

4 Conclusion

The proposed segmentation method is a promising approach for liver segmentation of contrast enhanced liver MRI. It is capable of segmenting liver organs with atypical shape as shown in cases from clinical routine. To date, it remains a tendency towards underestimation of the volume, and in few cases inclusion of the kidney into the segmented area is observed. Further evaluation with more cases and different MRI scanner manufacturers are envisioned to enhance robustness and accuracy.

References

1. Campadelli P, Casiraghi E, Esposito A. Liver segmentation from computed tomography scans: A survey and a new algorithm. *Artif Intell Med.* 2009;45:185–96.
2. Heimann T, van Ginneken B, Styner MA, et al. Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Trans Med Imaging.* 2009 Aug;28:1251–65.
3. Huppertz A, Balzer T, Blakeborough A, et al. Improved detection of focal liver lesions at mr imaging: multicenter comparison of gadoxetic acid enhanced mr images with intraoperative findings. *Radiol.* 2004;230:266–75.
4. Logeswaran R, Haw TW, Sarker SZ. Liver isolation in abdominal MRI. *J Med Syst.* 2008;32:259–68.
5. Platero C, González P, Tobar MC, et al. Automatic method to segment the liver on multi-phase MRI. In: *CARS*; 2008.
6. Massoptier L, Casciaro S. Fully automatic liver segmentation through graph-cut technique. *Proc IEEE Eng Med Biol Soc.* 2007;2007:5243–6.
7. Park H, Bland P, Meyer C. Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE Trans Med Imaging.* 2003;22(4):483–92.
8. Lim SJ, Joeng YY, Ho YS. Segmentation of the liver using deformable contour method on CT images. *Lect Notes Computer Sci.* 2004;3767:570–81.

Video-basiertes Tracking eines Bronchoskops

Design und quantitative Evaluierung

Tobias Reichl¹, Oliver Kutter¹, Benedikt Schultis¹, Manuela Menzel²,
Hubert Hautmann², Nassir Navab¹

¹Computer-Aided Medical Procedures (CAMP), TUM, München

²Medizinische Klinik I, Klinikum rechts der Isar, TUM, München

reichl@in.tum.de

Kurzfassung. Bronchovideoskopie ist eine weit verbreitete Technik für Diagnostik und Therapie von Atemwegserkrankungen. Aufgrund der Selbstähnlichkeit der Bronchien und der Bewegung des Bronchoskops ist die Orientierung erschwert, so dass eine Hilfestellung für Ärzte in Form von Navigationssystemen für die Bronchoskopie dringend notwendig ist. Wir führen mit Hilfe von sog. virtueller Bronchoskopie eine intensitätsbasierte Registrierung einer Sequenz von zweidimensionalen Videobildern mit einem dreidimensionalen CT-Datensatz auf der Grafikkarte durch. Dies ermöglicht für jedes Videobild in der Sequenz eine Positionsbestimmung der Kamera innerhalb des CT-Datensatzes. Wir stellen unsere Implementierung vor und schätzen insbesondere die erreichte Genauigkeit über die Präzision des Verfahrens ab.

1 Einleitung

Bronchovideoskopie ist eine weit verbreitete Technik für Diagnostik und Therapie von Atemwegserkrankungen.

Üblicherweise müssen die durchführenden Ärzte sich dabei auf ihr Wissen über anatomische Strukturen und das Videobild des Bronchoskops verlassen. Durch die Selbstähnlichkeit der Bronchien wird eine Orientierung allerdings erschwert, so dass eine Hilfestellung in Form von Navigationssystemen für die Bronchoskopie dringend notwendig ist. Hierbei kann einerseits die aktuelle Position und Orientierung des Bronchoskops im Verhältnis zur Anatomie des Patienten mit geeigneten Mitteln visualisiert werden, andererseits kann eine solche Positionsbestimmung auch zur Darstellung zusätzlicher Informationen benutzt werden. Beispielsweise kann die Lage von sonst nicht sichtbaren Strukturen wie Lymphknoten, Tumoren oder Gefäßen unter der Atemwegs Oberfläche eingeblendet werden.

Das grundlegende Problem für die Positionsbestimmung ist die automatische Registrierung von zweidimensionalem Videobild mit einem dreidimensionalen CT-Datensatz. Intensitätsbasierte Registrierung mit Hilfe von sog. virtueller Bronchoskopie wurde bereits von Mori et al. [1, 2] vorgeschlagen und bleibt vorerst die am meisten versprechende Möglichkeit zur Positionsbestimmung anhand von Videosequenzen. Insbesondere sind im Bereich der Atemwege doch deutlich

weniger optisch markante Punkte als Landmarken verfügbar als beispielsweise in der Speiseröhre.

Im Folgenden stellen wir unsere Implementierung und einen Ansatz zur quantitativen Abschätzung der erreichten Genauigkeit vor.

2 Materialien und Methoden

Für die Kalibrierung der intrinsischen Parameter der Bronchoskop-Kamera verwendeten wir die Methode von Zhang et al. [3], für die Entzerrung des Kamerabilds wurde die Methode von Heikkilä und Silven [4] benutzt.

Zur Darstellung der Atemwege verwenden wir ein Iso-Oberflächen-Rendering, ein Anzeigebeispiel ist in Abb. 1 ersichtlich. Ausgehend von einer Startpose mit ungefähr bekannter Pose (eine solche kann intraoperativ beispielsweise durch den Klick auf ein Pedal festgelegt werden), wird iterativ zu jedem Frame dasjenige virtuelle Bild gesucht, das die höchste Ähnlichkeit mit dem echten Bild aufweist. Eine solche Optimierung der extrinsischen Parameter der Kamera ergibt pro Frame eine Position und Orientierung für die Kamera innerhalb des CT-Datensatzes. Zusätzlich gewährleisten wir bei jeder Evaluierung der Kostenfunktion, dass der Optimierer das Innere der Bronchien nicht verlässt.

Für die Optimierung wurde der Simplex-Optimierer verwendet, als Ähnlichkeitsmaß haben wir Normalized Cross Correlation (NCC) verwendet. Sowohl die Erzeugung der virtuellen Bronchoskopie-Bilder als auch die Berechnung des Ähnlichkeitsmaßes fanden beide innerhalb eines eigenen Frameworks mit Hilfe von OpenGL/GLSL auf der Grafikkarte (GPU) statt.

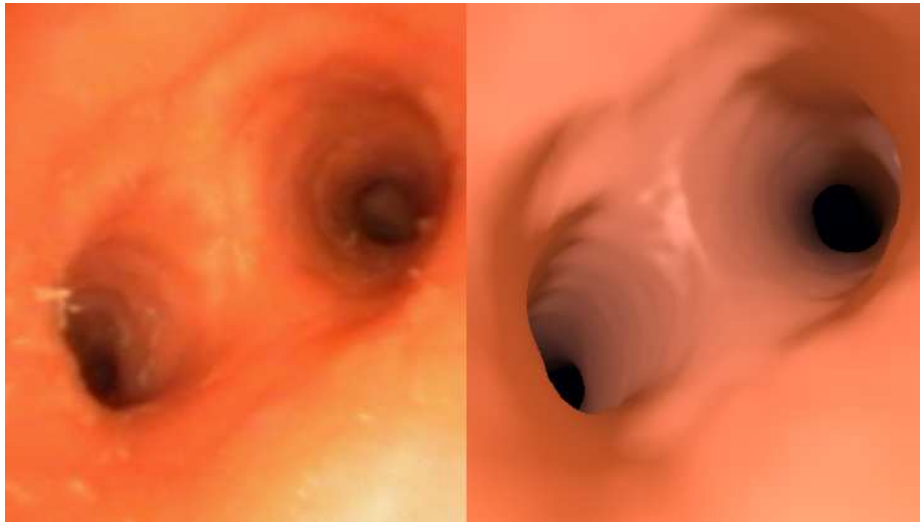


Abb. 1. Anzeigebeispiel mit realem Bronchoskopiebild (links) und entsprechendem virtuellem Bronchoskopiebild (rechts).

Um die Genauigkeit quantitativ abschätzen zu können, verwenden wir die statistische Verteilung der berechneten Positionen über mehrere Durchläufe hinweg, ähnlich dem Vorgehen von Atmosukarto et al. [5], d.h. wir schätzen die Genauigkeit anhand der Präzision (Wiederholungs-Genauigkeit) ab. Um nicht-deterministisch eine Mehrzahl von Durchläufen über dieselbe Videosequenz aufzeichnen zu können, addieren wir nach der Optimierung der Pose in jedem Frame Rauschen mit wählbarer Standardabweichung zur optimierten Pose hinzu. Über die Zeit hinweg akkumulieren die verschiedenen Durchläufe jeweils Rauschen, andererseits konvergieren die Durchläufe in jedem Frame idealerweise anhand der in den Bildern enthaltenen Informationen durch die Optimierung auf dieselbe Position. Die Standardabweichung dieser verschiedenen Positionen pro Frame ergibt eine untere Schranke für den Root Mean Square (RMS) Fehler für das Tracking. Qualitativ lässt sich aus einer Unempfindlichkeit gegenüber verrauschten Startpositionen für die Optimierung auch eine Robustheit gegenüber Störungen beim Tracking ableiten.

Die Experimente wurden mit dem CT-Datensatz eines Phantoms eines Bronchialbaums durchgeführt ($512 \times 512 \times 141$ Voxel, Auflösung $0.47 \times 0.47 \times 2$ mm). Das Video wurde mittels eines BF-1T180 Bronchovideoskops (Olympus Deutschland) von einem erfahrenen Pneumologen aufgezeichnet. Das von uns hinzugefügte Rauschen war normalverteilt mit Standardabweichungen zwischen 0.0625 und 1 Millimeter, jedes Experiment bestand aus zehn Durchläufen über die Videosequenz.

3 Ergebnisse

Für jeweils 95 % aller Frames war der Standardfehler (σ/\sqrt{n}) der Abweichung von der mittleren Position kleiner als 0.52 Millimeter. Wie in Abb. 2 ersichtlich ist die Präzision (und damit die untere Schranke der Tracking-Genauigkeit) über einen Großteil der Videosequenz hinweg besser als zwei Millimeter, später besser als drei Millimeter, wohlgemerkt ohne zusätzliche Informationen z.B. durch elektromagnetisches Tracking oder ähnliches.

4 Diskussion

Eine Evaluierung der Tracking-Genauigkeit ist grundsätzlich schwierig, da die zugrunde liegenden wahren Positionen meist nicht bekannt sind, sofern Positionen nicht beispielsweise über Fluoroskopie [6] oder CT-Aufnahmen [7] bestimmt werden. Für rein video-basierte Verfahren werden in der entsprechenden Literatur meist nur Angaben darüber gemacht, über welchen Zeitraum die Position verfolgt werden konnte, dies ist aber natürlich sehr stark vom verwendeten Datensatz abhängig. Qualitative Angaben fehlen zumeist und machen eine Beurteilung schwierig. Insbesondere für den direkten Vergleich verschiedener Trackingverfahren sind zudem quantitative Bewertungen wünschenswert.

Wie zu erwarten war, ist die Tracking-Präzision deutlich besser in Bereichen mit ausgeprägten Merkmalen wie beispielsweise Bifurkationen (Abb. 1) und deutlich schlechter in Abschnitten ohne derartige Merkmale.

Ein Vorteil des videobasierten Tracking ist, dass es stets eine anatomische Position innerhalb des statischen CT-Datensatzes liefert, d.h. unabhängig von Lageänderungen oder Atembewegungen des Patienten.

Mit dem vorliegenden System ist nur ein inkrementelles Tracking möglich, dies kann aber durch Vorhersage der weiteren Bewegung beispielsweise durch Kalman-Filter abgemildert werden [8, 9], ebenso können logische Informationen über die lokale Anatomie verwendet werden [7]. Zusätzlich ist es sinnvoll, eine Möglichkeit zur Reinitialisierung zu schaffen, zum Beispiel durch die Verwendung von elektromagnetischem Tracking [10, 2], das stets eine ungefähre Positionsangabe liefern kann.

Insbesondere stellen wir eine Möglichkeit vor, die Genauigkeit des Trackings quantitativ abzuschätzen.

Danksagung. Diese Arbeit wurde von der Deutschen Forschungsgemeinschaft im Rahmen des Projekts “Navigierte Bronchoskopie” (NA 620/2-1) gefördert.

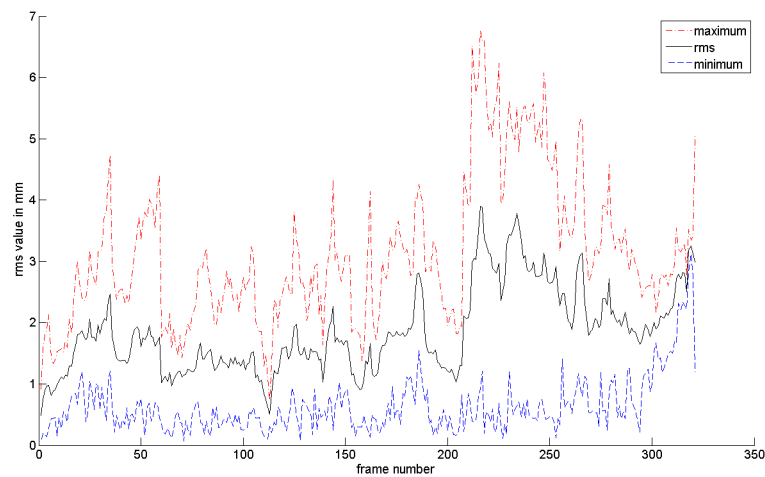


Abb. 2. Verlauf der Standardabweichung über die Videosequenz. Die Standardabweichung pro Frame variiert mit der Größe des hinzugefügten Rauschens, es wurden je Minimum, Durchschnitt und Maximum angegeben.

Literaturverzeichnis

1. Mori K, Suenaga Y, ichiro Toriwaki J, et al. Method for tracking camera motion of real endoscope by using virtual endoscopy system. *Proc SPIE*. 2000;3978:122–33.
2. Deguchi D, Ishitani K, Kitasaka T, et al. A method for bronchoscope tracking using position sensor without fiducial markers. *Proc SPIE*. 2007;6511:0N–1–9.
3. Zhang Z. A flexible new technique for camera calibration. *IEEE Trans Pattern Anal Mach Intell*. 2000;22(11):1330–4.
4. Heikkilä J, Silvén O. A four-step camera calibration procedure with implicit image correction. In: *Proc CVPR*; 1997. p. 1106–12.
5. Atmosukarto I, Soper TD, Glenn RW, et al. An interactive 3D user interface for guided bronchoscopy. *Proc SPIE*. 2007;6509:1G–1–13.
6. Hautmann H, Schneider A, Pinkau T, et al. Electromagnetic catheter navigation during bronchoscopy: validation of a novel method by conventional fluoroscopy. *Chest*. 2005;128(1):382–7.
7. Wegner I, Biederer J, Tetzlaff R, et al. Evaluation and extension of a navigation system for bronchoscopy inside human lungs. *Proc SPIE*. 2007;6509:1H–1–9.
8. Nagao J, Mori K, Enjouji T, et al. Fast and accurate bronchoscope tracking using image registration and motion prediction. *Lect Notes Comput Sci*. 2004;3217:551–8.
9. Mori K, Deguchi D, Kitasaka T, et al. Bronchoscope tracking based on image registration using multiple initial starting points estimated by motion prediction. *Lect Notes Comput Sci*. 2006;4191:645–52.
10. Mori K, Deguchi D, Akiyama K, et al. Hybrid bronchoscope tracking using a magnetic tracking sensor and image registration. *Lect Notes Comput Sci*. 2005;3750:543–50.

Image-based Quantification of Skin Irritation by Spatial Biomarker Profiling

Thora Pommerencke, Kathi Westphal, Claudia Ernst, Hartmut Dickhaus,
Niels Grabe

Institute for Medical Biometry and Informatics, University Hospital Heidelberg
Im Neuenheimer Feld 305, 69120 Heidelberg, Germany
thora.pommerencke@bioquant.uni-heidelberg.de

Abstract. We developed a method for the quantification of skin irritation based on the treatment of skin cultures and the subsequent automatic image analysis of fluorescently stained tissue sections being scanned with a fullslide scanner. In this method, the irritative effect is reflected in a shift of the expression pattern of selected biomarkers inside the epithelium. Using the example of SDS as irritant and HSP27 as a marker for skin irritation we showed that our method is capable of quantifying the irritative effect of SDS and furthermore delivering spatial and time dependent information.

1 Introduction

Every chemical or finished product has to be tested for its skin corrosive and irritation potential before it is made available to the human population. The prevailing risk assessment is conducted by means of *in vivo* animal tests. However, scientific concerns about the variability and predictive capacities of animal tests for humans and especially recent ethical reasons led to substantial efforts to develop alternative *in vitro* test methods. The commonly applied alternative method is the MTT test using organotypic culture systems, where the cell vitality is measured inside a homogenate [1]. This is suitable for testing skin corrosion, but not for mild irritation, as the latter does not lead to tissue destruction. In literature, these weak effects are often described qualitatively based on the visual inspection of histological stainings. We now developed a method for the quantification of skin irritation based on the automatic image analysis of intact, fluorescently stained tissue sections from treated skin cultures being scanned with a fullslide scanner. The irritative effect is reflected in a shift of the expression pattern of selected biomarkers inside the epithelium. We here used as irritant the detergent sodium lauryl sulphate (SDS) and the heat shock protein HSP27 [2] as a marker for skin irritation.

2 Methods

2.1 Experimental Setup

Mattek EFT 400 cultures were treated either with nothing, PBS or 0.4% SDS. For this, 25 μl of the respective solution was applied on top of the cultures on a mesh for different durations i.e. 1 h, 6 h, 16 h and 24 h. Afterwards, the cultures were cryo conserved and cut into sections of 6 μm . For each mode of treatment (e.g. 6 h SDS) there were four cultures and for each culture about 5 sections. The sections underwent an immunohistological fluorescent triple staining, labeling the cell nuclei with the nuclear dye DAPI in blue, the basal lamina using an antibody against the membrane protein Laminin 5 in green (Alexa 488) and the biomarker (i.e. Hsp27) for the quantification of skin irritation was marked in red (Alexa 594). The challenge was here to find a reference staining for the automatic segmentation of the epithelium which is stable against the SDS treatment. DAPI and Laminin proofed to be appropriate.

2.2 Image Acquisition

The fluorescent sections were scanned with the Nanozoomer HT from Hamamatsu Photonics at 20 \times magnification (a resolution of 0.46 $\mu\text{m}/\text{pixel}$). The images were recorded in three z-layers with a spacing of 2 μm each. The resulting slide scans were in a proprietary image format (.ndpi) from Hamamatsu. For further analysis those image regions to be analysed were selected in a computer assisted way and converted to be available for later image processing algorithms.

2.3 Segmentation of the Epithelium

As there was no staining marking the whole epithelium, the epithelial area had to be approximated using the staining of the basal lamina and the cell nuclei. However, having artificial skin cultures one can presume that the epithelium shows almost the same thickness (distance from basal lamina to surface) over the whole section neglecting border regions.

The first step was the segmentation of the basal lamina, visible in the green channel. After contrast enhancement and noise reduction mainly a thresholding based on the Otsu's method is applied. Since sometimes there was an unspecific

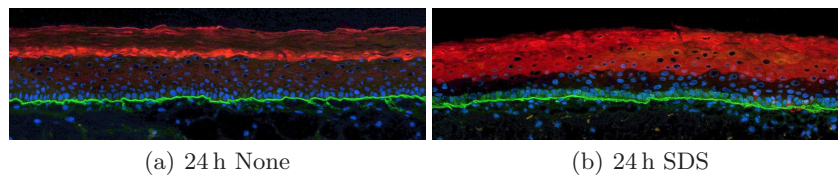


Fig. 1. Immunohistological fluorescent triple staining. Changed Hsp27 expression pattern following SDS treatment.

and diffuse staining in the stratum granulosum and corneum, the segmentation process is supported by an edge detection based on the Canny method parameterized to detect long stretched edges.

The cell nuclei were detected in the blue image channel by a two-step thresholding process [3], where only image areas close to strong image gradients were considered. Cell clusters were identified by size and split using a watershed method with local maxima as seeds. Finally, false positive cell nuclei were filtered out on the basis of object features like size, roundness and contrast. Ring-like stained nuclei, identified by the ratio of border intensity and inner intensity, were treated separately.

For the approximation of the surface line the basal lamina line was smoothed strongly and shifted towards the epithelium by 1.4 times the maximum (95 % quantile) nuclei distance from basal lamina. The shifting direction was determined by the nuclei density in a band around the basal lamina. The epithelial side exhibits a higher nuclei density than the side of the culture's collagen matrix.

To yield in a closed epithelial area with sides vertical to the surface, the surface pixels were fitted with a line and the endpoints of the surface were connected to the intersection points of the perpendicular through the respective endpoint and a point on the basal line.

2.4 Biomarker Profiling

For the biomarker profile the mean intensity in the red image channel was measured inside bands (given by distance intervals) of the determined epithelial area. The determined marker intensity was plotted against the corresponding "relative distance" [4] of the epithelial band. All profiles from one mode of treatment (e.g. 6h SDS) were averaged.

2.5 Feature Selection for the Quantification of Skin Irritation

In case of the protein Hsp27 SDS treatment leads to a premature and cytoplasmatic expression pattern with planar appearance (Fig. 1). In the profile a premature expression pattern is reflected by an early increase of the marker intensity.

To determine this starting point of expression, the section of the profile with the highest slope was fitted with a line. The intersection point of that line with the horizontal through the minimum marker intensity was set as the starting point of expression (*StartExpression*, Fig. 2a).

The planar histological expression pattern is reflected in a high integral of the profile from starting point of expression to profile's end. This feature *ProfileExpressionArea* was calculated by summing up all marker intensities from starting point of expression (Fig. 2b).

As the main change in expression can be observed between 40 % and 70 % relative distance also the sum of marker intensities between 40 % and 70 % relative distance was calculated (*ProfileExpressionArea40-70*). Furthermore, literature

[2] reports a translocation of Hsp27 into nuclei following SDS treatment. For this reason also the mean marker intensity located in the nuclei was measured in the area between 40 % and 70 % relative distance (*NucleiIntensity40-70*). All algorithms described above were implemented in Matlab R2008b.

3 Results

In Fig. 3 the median profiles with their 25 % and 75 % quantiles of two modes of treatment (24 h None and 24 h SDS) are shown exemplarily.

Setting the values of the extracted SDS features into relation with the corresponding None values, the SDS results show strong changes in all of the four features, strongest in ProfileExpressionArea40-70. At a treatment duration of 24 h the SDS value increases to a 2.6-fold of the corresponding None value. However, also a treatment of the cultures with PBS leads to a 1.5-fold increase in ProfileExpressionArea40-70. This means that even PBS has a somewhat irritative effect, possibly due to occlusive effects. Therefore, to measure the irritative effect of SDS, the SDS values are set into relation to the PBS values (Fig. 3c). At a treatment duration of 24 h a 1.7-fold increase in ProfileExpressionArea40-70 is detected.

4 Discussion

We have developed an automatic method for the quantification and spatio-temporal description of the effect of skin irritation inside the epithelium. Even though the tissue is morphologically still quite intact (not corroded yet), the irritative effect of SDS and even PBS can be measured within a good confidence interval (Fig. 3(a, b)). Complemented by further biomarkers this system could give important insights into the mechanisms of skin irritation. Having also measured an effect at PBS treatment makes this method promising for the test of mildly irritating substances.

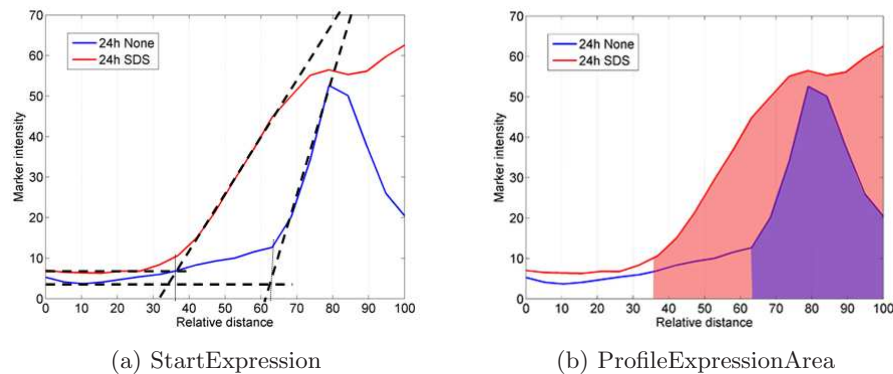
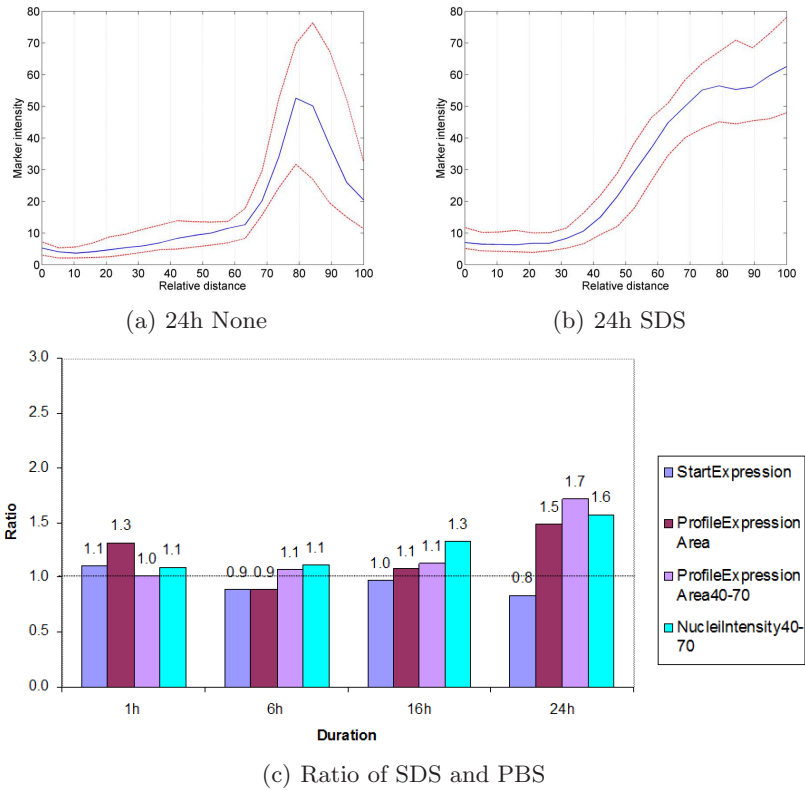


Fig. 2. Feature extraction from the marker profiles.

Fig. 3. Exemplary marker profiles with confidence interval (a,b). Time resolved change in expression after SDS treatment compared to PBS (c).



In future, further reagents like those mildly irritating (e.g. CTAB) have to be tested. To optimize the automatic segmentation, efforts should be put on establishing other histological reference markers allowing for a segmentation, which is independent of epithelial thickness.

References

1. Mosmann T. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J Immunol Methods*. 1983;65(1-2):55–63.
2. Boxman ILA, Hensbergen PJ, van der Schors RC, et al. Proteomic analysis of skin irritation reveals the induction of HSP27 by sodium lauryl sulphate in human skin. *Br J Dermatol*. 2002;146(5):777–85.
3. Pommerenke T, Dickhaus H, Grabe N. Vollautomatische einzelzellerkennung auf fluoreszenten gewebschnitten humaner epidermis. *Proc BVM*. 2009; p. 311–5.
4. Pommerenke T, Steinberg T, Dickhaus H, et al. Nuclear staining and relative distance for quantifying epidermal differentiation in biomarker expression profiling. *Bioinform*. 2008;9:473.

Interaktive Korrektur von 3D-Segmentierungen

Daniel Proksch, Jana Dornheim, Bernhard Preim

Institut für Simulation und Graphik, Universität Magdeburg
preim@isg.cs.uni-magdeburg.de

Kurzfassung. In dieser Arbeit werden vier neuartige Ansätze zur interaktiven Korrektur medizinischer 3D-Segmentierungen vorgestellt. Das Hauptziel besteht dabei darin, einen intuitiven sowie zeitsparenden Korrekturvorgang zu ermöglichen. Anstatt die segmentierten Objektkonturen in jeder einzelnen Schicht der tomographischen Datensätze voneinander getrennt zu bearbeiten, wird die Korrektur von uns als direkte Manipulation dreidimensionaler polygonaler Oberflächenmodelle interpretiert. Die vorgestellten Verfahren sind inspiriert von aktuellen Arbeiten zur benutzerfreundlichen Polygonmodellierung.

1 Einleitung

In vielen Fällen stimmen die Ergebnisse modellbasierter Segmentierungsverfahren nicht vollständig mit den gesuchten Objektkonturen überein. Ursächlich dafür sind nach [1] vor allem lokal schwache Bildinformation und Abweichungen der zugrundeliegenden Modellannahmen. Gerade bei der Verarbeitung medizinischer Bilddaten werden jedoch exakte Segmentierungen benötigt. Daher sind stets die Überprüfung der Ergebnisse durch erfahrene Mediziner und ggf. lokal manuelle Korrekturen erforderlich. Bei der Verwendung deformierbarer Modelle etwa, wird hierfür in der Regel aus der Voxelrepräsentation der Segmentierung eine editierbare Kontur generiert. Die Korrektur wird dann durch das einfache Überzeichnen der fehlerhaften Objektkontur in allen betroffenen (axialen) Schichten des tomographischen Datensatzes vorgenommen. So kann der Nutzer sein gesamtes Expertenwissen in den Korrekturvorgang einbringen. Nachteilig sind jedoch der hohe Zeitaufwand (insbesondere bei hochaufgelösten Datensätzen und ausgedehnten Objekten) sowie der stark subjektive Einfluss auf die Korrekturergebnisse. In dieser Arbeit werden effektive und intuitive Methoden zur Segmentierungskorrektur vorgestellt, die eine bessere Reproduzierbarkeit erwarten lassen.

2 Material und Methoden

Das bisherige Vorgehen zur Korrektur von Segmentierungen auf Basis deformierbarer Modelle ist vor allem deshalb ineffizient, weil bereits vorliegende Modellinformation vernachlässigt wird. Für den von uns präsentierten Ansatz wird daher nicht nur der Segmentierungsvorgang, sondern auch die Korrektur der daraus

hervorgehenden Ergebnisse als Deformierung dreidimensionaler Polygonmodelle interpretiert. Korrekturen sind somit nicht mehr nur auf einzelne Schichten begrenzt, sondern beeinflussen einen gewissen Nachbarschaftsbereich, wodurch der benötigte Zeitaufwand reduziert werden kann. Zudem ist es so möglich, beliebige Schnittführungen des Datensatzes zur Korrektur heranzuziehen.

Neben der Forderung nach geringem Zeitaufwand und Reproduzierbarkeit ist es nach [1] besonders wichtig, dass Korrekturwerkzeuge durch direkte Manipulation nachvollziehbar und zielgerichtet eingesetzt werden können. Inspiriert von Methoden zur intuitiven Polygonmodellierung für Anfänger und Gelegenheitsnutzer [2, 3], wurden von uns Prototypen für vier unterschiedliche Werkzeuge zur Korrektur medizinischer Segmentierungen entwickelt. Alle Interaktionen sind dabei als direkte Manipulation der Objektkonturen konzipiert, die sich als Schnitt des Segmentierungsmodells mit den zweidimensionalen Schichten des Datensatzes ergeben. Als Visualisierung stehen dem Nutzer axiale, saggitale und koronare Ansichten des Datensatzes sowie eine beliebig rotierbare, dreidimensionale Darstellung des Modells zur Verfügung.

2.1 Bulge Tool

Die Grundidee des Bulge Tools besteht darin, ein kugelförmiges Objekt durch den Datensatz zu bewegen, mit dem die Modellkontur gewissermaßen „ausgebault“ werden kann (Abb. 1a). Dabei hat der Nutzer zusätzlich die Möglichkeit, den Einflussbereich der Manipulation als Kugelradius zu definieren.

Abhängig von Größe und Auflösung des Modells steigt der Aufwand für die Detektion der Kollisionen von Kugel und Modellflächen, wodurch der flüssige Ablauf der Interaktion stark beeinträchtigt werden kann. Aus diesem Grund wird die Kollisionsdetektion auf die Schnittberechnung des kreisförmigen Abbildes der Kugel mit den Modellkonturen im zweidimensionalen Schichtbild reduziert. Liegt ein Konturbereich innerhalb des Kugelradius wird die Projektion des Kugelmittelpunktes auf die betreffenden Liniensegmente berechnet. Werden die so bestimmten Punkte in baryzentrischen Koordinaten bezüglich der entsprechenden Dreiecke dargestellt, lässt sich daraus eine sinnvolle Gewichtung für die Verschiebung der Dreiecksvertices ableiten. Schließlich müssen auch solche Vertices v berücksichtigt werden, die keiner der betrachteten Konturflächen angehören, sich aber dennoch innerhalb des Abstands r vom Kugelmittelpunkt m befinden. Die Verschiebung ergibt sich dann gemäß $v' = m + r \frac{v-m}{\|v-m\|}$.

2.2 Bending Tool

Das Bending Tool orientiert sich stark an der in [3] für die Animation von Charaktermodellen vorgestellten Technik zur Manipulation von Polygonmodellen. Die Deformation wird hierbei durch Einzeichnen einer Referenz- und einer Zielkurve bewirkt (Abb. 1b). Erstere entspricht zumeist in etwa der Hauptskeletlinie des zu deformierenden Modellteils. Anders als in [3] (wo vor allem annähernd zylindrische Objekte wie Arme und Beine transformiert werden), ist bei der

Anwendung zur Korrektur medizinischer Segmentierungen die explizite Begrenzung des Einflussbereiches (ROI) erforderlich. Dies geschieht durch Definition eines die Referenzkurve $C_r(s)$ umhüllenden Polygons sowie zweier parallel zum Schichtbild liegenden Grenzebenen. Das Einzeichnen der Zielkurve $C_t(s)$ führt schließlich zur Deformation.

Zunächst werden alle innerhalb der ROI liegenden Vertices v mit den jeweils nächstgelegenen Punkten v_r auf $C_r(s)$ sowie den korrespondierenden Punkten v_t auf $C_t(s)$ assoziiert. Für die Transformation wird dann jeweils v_r in den Koordinatenursprung verschoben, eine Rotation um die Schnittachse im Winkel θ zwischen den Kurventangenten in v_r bzw. v_t und schließlich eine Translation um v_t berechnet.

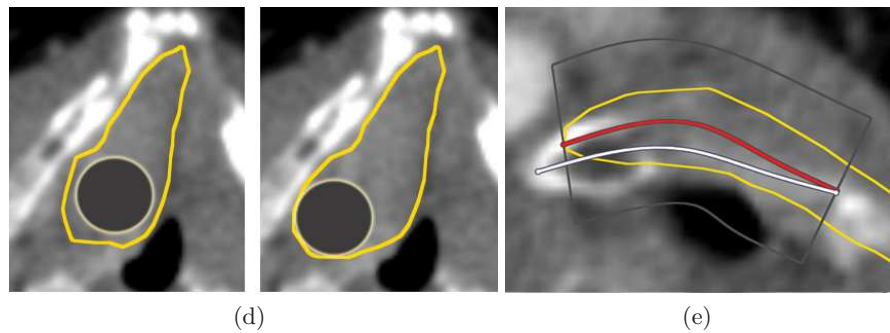


Abb. 1. (a) Bulge Tool: Mit einem kugelförmigen Werkzeug wird die Objektkontur „ausgebault“. (b) Bending Tool: Durch Einzeichnen einer Referenzkurve wird die Hauptskeletlinie angenähert. Eine zweite Kurve beschreibt das Deformationsziel. Durch das graue Hüllpolygon wird die ROI definiert.

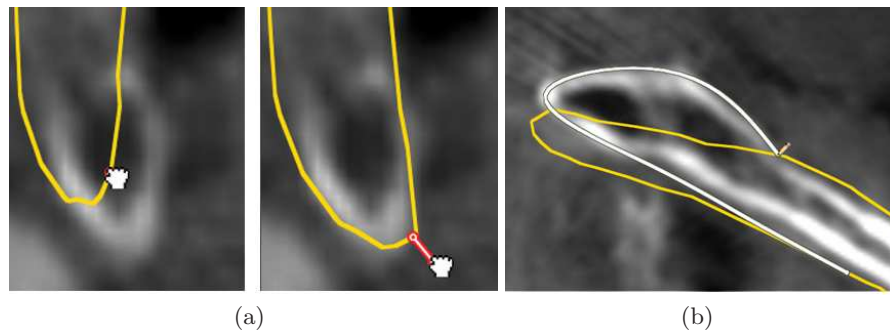


Abb. 2. (a) Traction Tool: Durch „Greifen“ und „Verziehen“ der Kontur wird das Modell unter Beibehaltung lokaler Formeigenschaften verformt. (b) Sketch Tool: Das Überzeichnen (weiß) der fehlerhaften Modellkontur resultiert in einer dreidimensionalen Deformation, ebenfalls unter Berücksichtigung von Formeigenschaften.

2.3 Traction Tool

Besonders in Fällen, in denen Fehlsegmentierungen auf Bildartefakten im Datensatz und nicht auf unzureichenden Modellannahmen beruhen, ist es sinnvoll das bereits vorliegende Formwissen auszunutzen und somit weniger subjektiv beeinflusste Endergebnisse zu erzielen.

Um dies zu erreichen, wird für das Traction Tool vom *Laplacian Modeling Framework* (LMS) [4] Gebrauch gemacht. Das LMS erlaubt es, Polygonmodelle unter Aufrechterhaltung lokaler Formeigenschaften zu manipulieren. Ausgehend von der Berechnung diskreter Laplace-Koordinaten entsteht ein lineares Gleichungssystem (LGS). Die Deformation lässt sich nun durch Hinzufügen positiver Bedingungen (als weitere Zeilen) beschreiben. Die näherungsweise Lösung des somit überbestimmten LGS führt zu einer optimalen Deformation unter Aufrechterhaltung lokaler Formeigenschaften sowie der Berücksichtigung aller weiteren Bedingungen.

Bei der Anwendung des Traction Tools wird zunächst ähnlich wie beim Bulge Tool der Radius der ROI bestimmt. Zur Deformierung klickt der Nutzer anschließend auf den zu korrigierenden Teil der Modellkontur und „zieht“ durch Bewegung der Maus daran (Abb. 2a). Intern werden dabei derart lineare Zusatzbedingungen in das LGS aufgenommen, so dass die Randvertices der ROI an ihrer Ursprungsposition verankert bleiben und dem ursprünglichen Ort des Mausclicks auf der Modelloberfläche die jeweils aktuelle Cursorposition zugewiesen wird.

2.4 Sketch Tool

Skizzenbasierte Polygonmodellierung wie in [2] ist eine für unerfahrene Nutzer konzipierte Technik, an der sich das Sketch Tool orientiert. Die Interaktion erfolgt als Kurvenzeichnung (Abb. 2b), so dass ähnlich wie bei der bisherigen manuellen Korrektur, fehlerhafte Bereiche der Segmentierung einfach überzeichnet werden. Im Unterschied dazu, bewirkt das Sketch Tool entsprechend unserem Ansatz jedoch eine dreidimensionale Modellmanipulation. Um vorhandenes Formwissen ausnutzen zu können, basiert auch das Sketch Tool auf dem LMF.

Die ROI ergibt sich hier durch zwei Ebenen, die jeweils senkrecht zum Start- und Endpunkt der gezeichneten Kurve stehen, sowie zwei manuell definierter Grenzebenen parallel zur Bildebene. Nach der Interaktion wird der überzeichnete Teil des fehlerhaften Segmentierungsmodells automatisch als jener Konturteil identifiziert, der die größte Ähnlichkeit mit der Nutzerkurve aufweist. Der so gefundene Teil der Kontur wird als parametrische Kurve interpretiert. So können neben der Verankerung der Randvertices solche Bedingungen in das LGS aufgenommen werden, die einer Menge von Punkten der ursprünglichen Kontur die korrespondierenden Positionen auf der Nutzerkurve zuordnen.

3 Ergebnisse

Die vorgestellten Interaktionstechniken wurden unter Verwendung der Entwicklungsumgebung *MeVisLab* sowie der *Magdeburg Shape Model Library* prototy-

pisch in C++ implementiert und in einer ersten qualitativen Nutzerstudie zur Korrektur modellbasierter Segmentierungen in Hals-CTs eingesetzt (eine quantitative Evaluierung unter Mitwirkung medizinisch geschulter Anwender steht bisher aus). Aufgrund mangelnder Fachkenntnisse hatten die fünf Teilnehmer die Aufgabe Segmentierungsergebnisse an von Medizinern erstellte Referenzsegmentierungen anzupassen. Hierbei zeigte sich die generelle Tauglichkeit aller vier Ansätze. Als besonders vielversprechend erwies sich das Traction Tool, da es insbesondere bei der Anwendung in verschiedenen Schnittführungen (einem wesentlichen Vorteil gegenüber der bisherigen Art der Korrektur) überzeugen konnte.

4 Diskussion

Die von uns vorgestellten Interaktionstechniken ermöglichen eine neuartige Herangehensweise für die Korrektur modellbasierter Segmentierungen. Auch darüber hinaus erscheint die Anwendung sinnvoll. Tumoren in Hals-CTs etwa, werden unter Zuhilfenahme endoskopischer Untersuchungsergebnisse in der Regel vollständig manuell segmentiert. Werden bereits vorhandene Segmentierungen in polygonale Oberflächenmodelle überführt, können (durch Tumorwachstum oder Therapie) erforderliche Modifikationen mit unserem Ansatz leicht realisiert werden. Darüber hinaus ist zu jedem Zeitpunkt eine dreidimensionale Visualisierung der Modelle verfügbar, so dass die Anpassung an zuvor gesehene Bilder erleichtert wird.

Der nächste wichtige Schritt für die Weiterentwicklung der Korrekturwerkzeuge besteht in der Durchführung einer ausgedehnten Evaluierung mit Fachpersonal. Dabei sollte vor allem untersucht werden, inwieweit sich ein Zeitgewinn gegenüber der bisherigen Herangehensweise erzielen lässt. Außerdem ist von Interesse, ob sich einzelne Werkzeuge durchsetzen oder die Testpersonen Kombinationen der verschiedenen Techniken bevorzugen.

Literaturverzeichnis

1. Olabarriaga SD, Smeulders AWM. Setting the mind for intelligent interactive segmentation: overview, requirements, and framework. In: Proc IPMI; 1997. p. 417–422.
2. Igarashi T, et al. Teddy: a sketching interface for 3D freeform design. In: Proc SIGGRAPH; 1999. p. 409–416.
3. Kho Y, Garl M. Sketching Mesh Deformations. In: Proc SIGGRAPH Symp I3D; 2005. p. 147–154.
4. Lipman Y, Sorkine O, Alexa M, et al. Laplacian framework for interactive mesh editing. Int J Shape Model. 2005; p. 43–61.

Graphen- und Level-Set-basierte Nachverarbeitung von 3D-Gefäßsegmentierungen

Nils Daniel Forkert¹, Alexander Schmidt-Richberg¹, Dennis Säring¹,
Jens Fiehler², Till Illies², Dietmar Möller³, Heinz Handels¹

¹Institut für Medizinische Informatik,

²Klinik und Poliklinik für Neuroradiologische Diagnostik und Intervention,
Universitätsklinikum Hamburg-Eppendorf, 20246 Hamburg

³Arbeitsbereich Technische Informatiksysteme, Universität Hamburg

`n.forkert@uke.uni-hamburg.de`

Kurzfassung. Die Ergebnisse von automatischen Verfahren zur Segmentierung von zerebralen Gefäßen benötigen in der Regel eine Nachverarbeitung. Ein häufig auftretendes Problem sind Unterbrechungen in den Gefäßsegmentierungen, meistens hervorgerufen durch geringe Intensitäten der Gefäße. Diese Unterbrechungen wirken sich häufig negativ auf nachfolgende bildgestützte Analysen aus. In diesem Beitrag wird ein automatisches Verfahren zur Korrektur von Unterbrechungen in Gefäßsegmentierungen vorgestellt. Hierbei wird zunächst auf Basis der Gefäßsegmentierung das 3D-Skelett berechnet, welches im Folgenden dafür verwendet wird, Gefäßendpunkte zu detektieren. Auf Basis des Vesselness-Bildes wird ein kantengewichteter Graph konstruiert, der zur Detektion möglicher Verbindungen zwischen den Gefäßendpunkten verwendet wird. Nach einer Konsistenzprüfung werden übrig gebliebenen Pfade abschließend verwendet, um mittels eines Level-Set-Ansatzes die finale Segmentierung zu generieren. Zur Evaluation wurde ein synthetischer Datensatz sowie ein klinischer Datensatz verwendet. Die Ergebnisse zeigen, dass die vorgestellte Methode Unterbrechungen findet und korrigieren konnte. Durch die vorgestellte Methode kann die Genauigkeit von Gefäßsegmentierungen für nachfolgende Analysen deutlich verbessert werden.

1 Einleitung

Eine genaue Segmentierung von Gefäßen, basierend auf hochauflösenden 3D-Angiographie-Bildsequenzen ist für viele Anwendungen von hoher Bedeutung. So variiert z.B das Ergebnis von FEM-Blutfluss-Simulationen [1] zum Teil erheblich mit der Genauigkeit der Gefäßsegmentierung. Zur Gefäßextraktion aus medizinischen Bildsequenzen wurde eine Vielzahl an Methoden vorgestellt, für eine detaillierte Übersicht sei an dieser Stelle auf die Arbeit von Suri et al. [2] verwiesen. Ein Problem, dass sich bei vielen insbesondere intensitätsbasierten Ansätzen ergibt, sind Unterbrechungen in der finalen Segmentierung. Obwohl einige Segmentierungsansätze einen Nachverarbeitungsschritt integrieren (z.B. [3]), ist ein genereller Ansatz für dieses Problem den Autoren nicht bekannt. In diesem

Beitrag wird eine Methode zur automatischen Detektion und Verbindung von Unterbrechungen in Gefäßsegmentierungen vorgestellt.

2 Material und Methoden

2.1 Detektion von Gefäßenden

Zur Korrektur von Unterbrechungen in einer bestehenden Gefäßsegmentierung müssen zunächst die korrespondierenden Gefäßenden detektiert werden. Hierzu wird das 3D-Skelett von der Segmentierung berechnet. Im Rahmen dieser Studie wurde hierzu das Verfahren von Lee und Kashyap [4] verwendet. In dieser Darstellung lassen sich die Gefäßenden leicht über eine Nachbarschaftsanalyse bestimmen, wobei ein Gefäßende dadurch charakterisiert ist, dass ein Voxel des Skeletts genau einen Skelettvoxel als Nachbarn aufweist.

2.2 Graphenbasierte Verbindung der Gefäßenden

Um die detektierten Gefäßenden zu detektieren, wurde ein gerichteter Kostengraph $G(V, E)$ konstruiert, wobei die Knoten v_i aus der Knotenmenge V des Graphen die Voxel des Bildes definieren. Die Knoten von im Bildraum benachbarten Voxeln werden dabei mit Kanten aus der Kantenmenge E verbunden. Zur Zuweisung der Kantenkosten kann zum Beispiel das Vesselness-Parameterbild [5] verwendet werden. Der Vesselnessfilter hat hierbei den Vorteil insbesondere kleine Gefäße, die sich meist durch geringe Intensitäten im Originalbild auszeichnen und von intensitätsbasierten Segmentierungsverfahren nicht erkannt werden, gegenüber dem Hintergrund hervorzuheben. Da bereits segmentierte Gefäßvoxel für die nachfolgende Analyse nicht von Interesse sind, wird das Kostenbild mit der invertierten Gefäßsegmentierung maskiert. Das entstandene Bild wird nun mit einem Schwellwertfilter bearbeitet, so dass Werte mit einem höheren Wert als I auf I gesetzt werden. Einzelne sehr hohe Vesselnesswerte können in der nachfolgenden graphenbasierten Analyse zur Detektion von Pfaden durch nicht-vaskuläre Gewebe führen. Die Schwellwertfilterung dient dazu die Anzahl solcher Pfade zu reduzieren. Da das Problem Gefäßunterbrechungen zu finden nachfolgend als Kürzester-Pfad-Problem betrachtet wird, wird das verarbeitete Kostenbild abschließend noch invertiert und dann dazu verwendet die Kantenkosten des Graphen zu definieren. Das Problem, eine Verbindung zwischen den Knoten von korrespondierenden Gefäßenden im Graphen zu finden, kann als Kürzester-Pfad Problem beschrieben werden. Ein Pfad p ist dabei eine Folge von Knoten $\langle v_0, v_1, \dots, v_k \rangle$ im konstruierten Graphen $G(V, E)$, wobei die Kanten (v_i, v_{i+1}) in der Kantenmenge E vorkommen. Das Gewicht w eines Pfades ist dabei die Summe der Gewichte aller im Pfad vorkommenden Kanten

$$w(p) = \sum_{i=0}^{k-1} w(v_i, v_{i+1}) \quad (1)$$

Der kürzeste Pfad von einer Menge an möglichen Pfaden zwischen zwei Kanten v_0 und v_k ist definiert als der Pfad, der die geringsten Kosten aufweist. Zur Berechnung der kürzesten Pfade wurde in dieser Arbeit der Algorithmus von Dijkstra's verwendet. Ausgehend vom ersten gefunden Gefäßende können nun die minimalen Kosten vom korrespondierenden Knoten v_s im Graphen zu allen weiteren Gefäßenden v_{s+1}, \dots, v_n berechnet werden, wobei n die Anzahl der Gefäßenden darstellt. Die erhaltenen Kosten können mittels der Pfadlängen normiert werden, um so Pfade unabhängig von ihrer Pfadlänge gleichberechtigt zu analysieren. Falls die Pfadkosten unter einem definierten Schwellwert λ liegen, wird dieser Pfad als eine Gefäßverbindung definiert. Auf diese Weise werden ausgehend von einem Gefäßende mehrere Verbindungen erlaubt, was im Falle von Bifurkationen hilfreich ist.

2.3 Level-Set-Segmentierung

Um die extrahierten Pfade an die Gefäße anzupassen, wurde ein variationeller Level-Set-basierter Segmentierungsansatz verwendet [6]. Die optimale Level-Set-Funktion wird durch die Minimierung des Energiefunktionals

$$\mathcal{J}[\phi] := \mathcal{I}[\phi] + \mathcal{E}[V; \phi] \quad (2)$$

wobei $V(x)$ das Vesselness Parameterbild ist. Das Funktional besteht aus zwei Termen: der internen Energie \mathcal{I} , welche dazu dient die Grenzen glatt zu halten und einem Regionen-basierten äußeren Energieterm

$$\mathcal{E}[V; \phi] := \int_{\Omega} H(\phi(x)) \cdot \log p_{\text{out}}(V(x)) + (1 - H(\phi(x))) \cdot \log p_{\text{in}}(V(x)) dx \quad (3)$$

wobei H die Heavyside-Funktion bezeichnet, welche dazu verwendet wird, das Innere und Äußere des Objektes zu beschreiben. Mittels (3) kann a-priori Wissen über die Intensitätsverteilung innerhalb p_{in} und außerhalb p_{out} der Gefäße einbezogen werden. Basierend auf der vorhandenen Segmentierung und dem Vesselness-Parameterbild können diese Verteilungen mittels einer Parzen-Window Strategie abgeschätzt werden.

Für die Initialisierung wird auf Basis der detektierten Pfaden p zunächst ein Binärbild generiert. Eine anschließende Dilatation um ein Voxel dient dazu, den Verlust von dünnen Strukturen zu verhindern.

2.4 Experimente

Die vorgestellte Methode wurde zunächst anhand synthetischer Daten getestet, um die generelle Funktionsweise zu testen. Hierbei wurde ein synthetischer 3D-Datensatz einer Spirale generiert, wobei die Intensität einzelner Abschnitte schrittweise verringert worden ist. Da klinische Datensätze meistens Rauscharfakte beinhalten wurde der synthetische Datensatz mit Rauschen überlagert, um so realitätsnähere Verhältnisse zu schaffen. Abschließend wurden die hohen

Intensitäten des Datensatzes mittels Schwellwertfilter extrahiert. In einem weiteren Experiment wurde auf Basis eines realen 3D Time-of-Flight (TOF) MRA Datensatzes zunächst das Gefäßsystem mittels dem von Chapman et al. [7] vorgestellten Volume-Growing basierten Z-Buffer Segmentierung extrahiert und dann die vorgestellte Methode auf diese Segmentierung angewendet. Die Evaluation erfolgte dabei von medizinischen Experten, indem die gefundenen Verbindungen hinsichtlich der Plausibilität qualitativ untersucht wurden. Als Schwellwerte wurden bei beiden Experimenten die Parameter $I = 25$ und $\lambda = 4$ verwendet.

3 Ergebnisse

Abb. 1 zeigt die Ergebnisse der einzelnen Schritte des vorgestellten Verfahrens basierend auf dem generierten synthetischen Datensatz mit den Unterbrechungen mit unterschiedlichen Intensitäten. Hierbei zeigt sich, dass mittels der graphen und Level-Set-basierten Methode alle Unterbrechungen gefunden und korrigiert werden konnten.

Abb. 2 zeigt das Oberflächenmodell eines segmentierten TOF-Datensatzes, in dem sich die ursprüngliche Segmentierung rot und die neu hinzugekommenen Verbindungen türkis darstellen. Hierbei wurden 207 Gefäßenden und 45 Verbindungen gefunden. Die visuelle Inspektion durch medizinische Experten ergab, dass es sich bei allen der 45 gefundenen Verbindungen um korrekte Gefäßverbindungen handelt. Es zeigte sich jedoch auch, dass einige Gefäßverbindungen nicht detektiert werden konnten, wobei die Anzahl nicht eindeutig quantifizierbar war.

4 Diskussion

Im Rahmen dieser Arbeit wurde ein Verfahren zum Finden und Nachverarbeiten von Unterbrechungen von Gefäßsegmentierungen vorgestellt. Die qualitative Plausibilitätsprüfung ergab, dass nicht nur kurze Unterbrechungen korrigiert wurden, sondern auch sehr lange Gefäßabschnitte detektiert werden konnten.

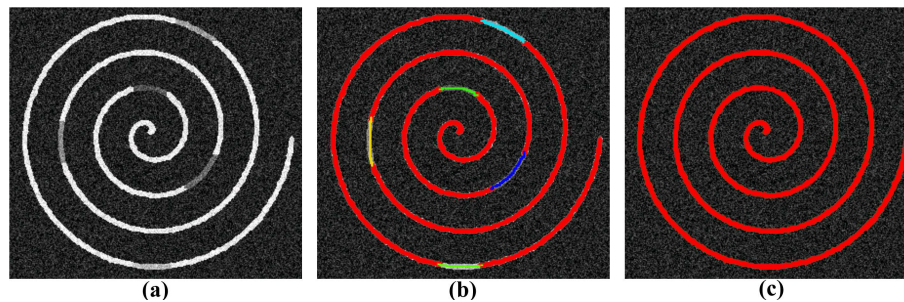
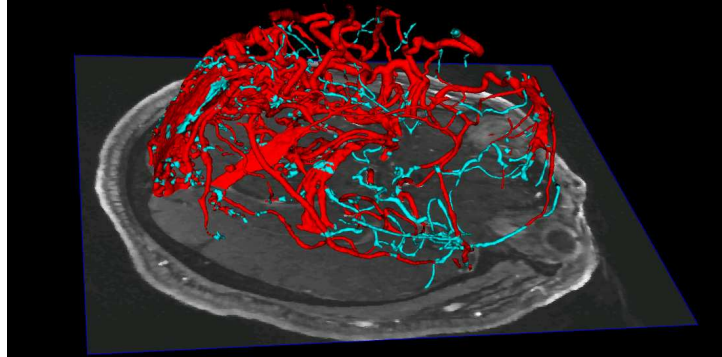


Abb. 1. Synthetischer 3D-Datensatz mit Abschnitten verringerter Intensität (a), Schwellwertbasierte Segmentierung (rot) und gefundene dilatierte Verbindungen von Gefäßenden (b) und finale korrigierte Segmentierung (c).

Abb. 2. 3D-Visualisierung des segmentierten Gefäßsystems: Initiale Z-Buffer Segmentierung (rot) und gefundene und korrigierte Gefäßunterbrechungen (türkis).



Zur weiteren Evaluation des vorgestellten Verfahrens sind weitere Datensätze notwendig. Zusätzlich ist eine Evaluation der verwendeten Parameter geplant, um die Ergebnisse der vorgestellten Methode zu optimieren. So könnte eine Verminderung der gewählten Schwellwerte dazu führen mehr Gefäßunterbrechungen zu korrigieren, jedoch auch zu falsch-positiven Ergebnissen.

Zusammenfassend kann das vorgestellte Verfahren die Qualität von Gefäßsegmentierungen und nachfolgende Analysen deutlich erhöhen.

Literaturverzeichnis

1. Jialiang C, Shengzhang W, Wei Y, et al. Computational fluid dynamics modeling of intracranial aneurysms. In: *Int Conf Biomed Eng Inform*; 2008. p. 566–9.
2. Suri J, Liu K, Reden L, et al. A review on MR vascular image processing algorithms: Skeleton versus nonskeleton approaches: Part II. *IEEE Trans Inf Technol Biomed.* 2002;6(4):338–50.
3. Hassouna MS, Farag A, Hushek S, et al. Cerebrovascular segmentation from TOF using stochastic models. *Med Image Anal.* 2006;10(1):2–18.
4. Lee T, Kashyap R, Chu C. Building skeleton models via 3D medial surface/axis thinning algorithms. *Graph Models Image Process.* 1994;56(6):462–78.
5. Sato Y, Nakajima S, Shiraga N, et al. Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Med Image Anal.* 1998;2(2):143–68.
6. Schmidt-Richberg A, Handels H, Ehrhardt J. Integrated segmentation and non-linear registration for organ segmentation and motion field estimation in 4D CT data. *Methods Inf Med.* 2009;48(4):344–9.
7. Chapman B, Stapelton J, Parker D. Intracranial vessel segmentation from time-of-flight MRA using pre-processing of the MIP Z-buffer: accuracy of the ZBS algorithm. *Med Image Anal.* 2004;8(2):113–26.

Robust Automatic Calcium Scoring for CT Coronary Angiography

Matthias Teßmann^{1,2}, Fernando Vega-Higuera², Bernhard Bischoff³,
Jörg Hausleiter³, Günther Greiner¹

¹Computer Graphics Group, University of Erlangen-Nürnberg, Germany

²Siemens AG, Sector Healthcare, Computed Tomography, Forchheim, Germany

³Klinik für Herz- und Kreislauferkrankungen, Deutsches Herzzentrum München

matthias.tessmann@informatik.uni-erlangen.de

Abstract. Currently, an increased number of studies are being performed in order to evaluate the clinical value of calcium scoring on contrast enhanced computed tomography coronary angiography images. One major finding is the increased diagnosis time and manual effort required to accurately segment calcified lesions in the contrast enhanced data. In this paper, a novel approach for the fully automatic segmentation and quantification of calcified lesions in coronary computed tomography angiography is presented. The algorithm includes a robust segmentation threshold determination based on an automatically generated vessel-tree histogram. Thereby, lesions can be accurately segmented and scores can be determined without user interaction. Validation against manual scores obtained by a radiologist yielded a very high correlation, which indicates the clinical value of the presented method.

1 Introduction

Cardiovascular events are currently the leading cause of death in developed countries. It is thus essential to identify high-risk patients for the prevention of secondary effects [1]. Since contrast enhanced CT coronary angiography (CTCA) has proven to allow a reliable diagnosis of cardiovascular stenoses [2], more and more clinical studies are published on the topic of establishing an equivalence between traditional coronary calcium (CAC) scoring and scoring methods on CTCA data [1, 3]. However, working with contrast enhanced data is a complex task requiring more time and manual effort than traditional methods [1]. In order to reduce the complexity and duration of the diagnosis process on CTCA data, there exist several algorithms that do provide semi-automatic detection of calcified lesions on this data, e.g. from Wesarg et al. [4]. However, these methods also require manual intervention to some degree. In this paper, a novel approach for fully automatic detection and quantification of calcified lesions on CT coronary angiograms is presented. Its major contributions are a robust estimation of a Hounsfield-Unit (HU) threshold for calcium detection and segmentation in contrast enhanced CT data as well as an accurate detection of lesions within

the vascular tree. The resulting algorithm has been validated against diagnostic data provided by radiological experts and showed to yield equivalent clinical results.

2 Materials and Methods

In order to examine a CTCA dataset for calcified coronary lesions, the entire vascular tree is extracted with the method presented in [5]. Having this tree is a prerequisite for the remaining two steps of the algorithm: HU threshold determination and lesion segmentation, as explained in the following sections.

2.1 HU-Threshold Determination

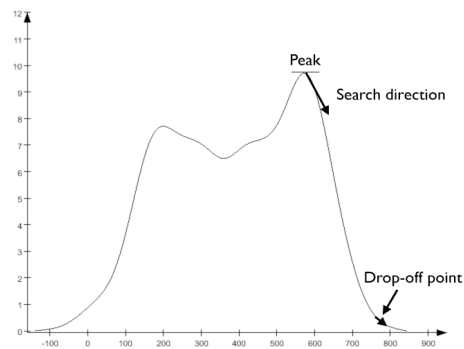
A major problem for the detection of an adequate calcium segmentation threshold is that image contrast in CTCA data has a high variability. This is due to varying contrast agent absorption rates over time and patient physiology. As a consequence, HU values for vessel lumen vary from patient to patient and from study to study. Therefore, a threshold has to be computed individually for each data set. For this purpose, the previously generated vascular tree can be exploited. In order to determine a threshold τ , a histogram $H(x)$ based on the intensity values of the voxels along the automatically segmented vessel centerlines is generated. It is assumed that by only taking into account the voxel values that are covered by the segmented vascular tree, the HU values for which the histogram function reaches its maximum are equivalent to the mean value of healthy vessel lumen. Moreover, it can be assumed that the number of voxels belonging to calcified plaques is not significant in comparison to the number of voxels belonging to contrast agent. Subsequently, the histogram function is scanned for the maximum number of hits. The width of the histogram area around this peak covers the distribution of vascular lumen within the histogram function. The ideal HU threshold for calcified plaque segmentation corresponds to the lowest intensity value that belongs to calcium but does not fall into the range of contrast agent. Hence, in order to find the optimal value for separation of vessel lumen and coronary calcium, it is required to find the point where the descent of the function starting at the peak position flattens. This can be done by computing the derivative of the histogram function $H'(x)$ and examining its values starting at the peak position until it drops below a certain threshold ϵ . Based on the ground truth database we found $\epsilon = -10^{-2}$ to be the optimal value. Selecting this point yields the desired threshold value τ (Fig. 1).

2.2 Detecting Calcifications

In the calcification detection step, the vascular tree is processed in a point-by-point fashion. At each centerline point the voxel neighborhood within a spherical radius of 1.5 mm is examined. If a voxel is found to have a HU value $> \tau$ it is marked as a possible calcium candidate (Fig. 2). The resulting candidate points

together with the threshold τ are then used as input seeds for a standard region growing algorithm as indicated by the clinically established calcium scoring methods [2, 6]. Next, the resulting segmentation mask is used as input to a connected components analysis in order to identify single lesions and remove lesions whose size is below or above a plausible value, e.g. 1-voxel lesions that may be detected due to image noise. Finally, the detected lesions are scored individually by the use of standard scoring algorithms: Agatston score [6], calcium volume and calcium mass on a per-lesion, per-vessel and overall calcium burden basis.

Fig. 1. Histogram of the centerline tree of a dataset from the test database. Marked are the peak value, descending direction and the drop-off point that is finally selected as segmentation threshold.



2.3 Validation Setup

For the validation of the presented algorithm 47 clinically relevant data sets were used. All calcified lesions contained in the data were marked manually by a

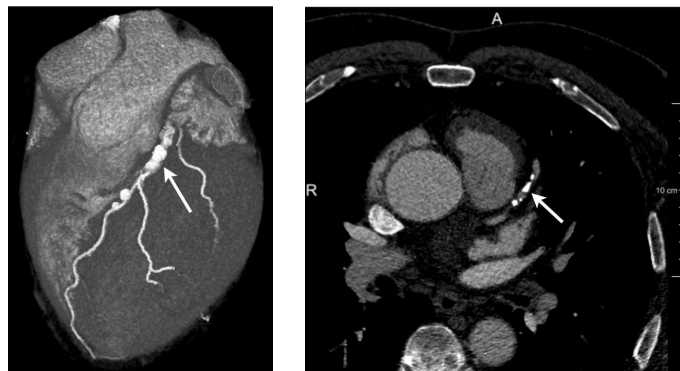


Fig. 2. Candidate selection result on an example dataset. Left: volume rendering of the heart with marked calcium positions (indicated by the arrow); right: corresponding slice image including calcified lesions.

Table 1. Correlation and limits of agreement between manual and automatic scoring.

Score Type	ρ	95% limits of agreement
Agatston Score	0.946	0.26 ± 29.78
Calcium Mass	0.951	0.41 ± 10.83
Calcium Volume	0.950	0.91 ± 22.85

radiologist. Each data set was assigned an individual HU threshold for a region-growing based segmentation of the coronary calcium. The resulting segmentation masks were used to generate ground-truth reference calcium scores (Agatston score, calcium mass score and calcium volume score) for evaluation. Finally, our method has been applied in unattended batch-mode on the same database. The resulting automatically detected HU thresholds and calcium scores were then compared to the ground truth data.

3 Results

The generated automatic thresholds are very close to those manually selected. The mean difference between the two values was found to be about 25 HU units. In 62% of the examined cases, the automatically detected segmentation threshold was lower than the segmentation threshold selected by the radiologist. However, no flooding occurred during region growing, which indicates a better threshold selection in comparison to the manually selected value. The correlation between manually and automatically generated scores expressed as Spearmans ρ and their respective 95% intervals of agreement [7] for the three scoring types are very high in all cases, as shown in Tab. 1. Plots comparing volume and mass scoring results directly are shown in Figure 3.

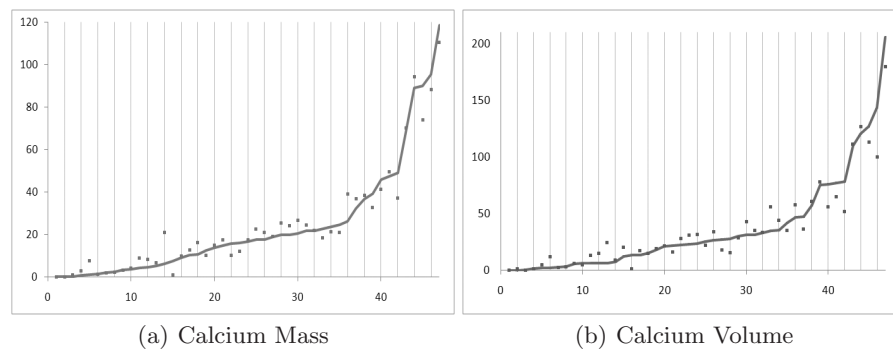


Fig. 3. Comparison of manual and automatically generated scores on the test database (Patient number and total score on the x- and y-axis, respectively). The line corresponds to the manual scoring result.

4 Discussion

The presented automatic calcium detection and segmentation method was shown to be adequate for calcium scoring on CTCA data when compared to results obtained by a radiological expert. Considering the broad cluster sizes for risk-grouping of patients in current clinical practice [8], the deviation of values with respect to the expert's ground truth is not significant. Additionally, most of the deviations from the reference scores are caused by the automatically determined segmentation threshold providing a better fit to the data than the manually selected threshold. Therefore, more voxels belonging to a calcified lesion are segmented, yielding an increased score value. It is reasonable to assume that the resulting higher score more accurately resembles the real calcium burden. In the immediate future, we would like to investigate the influence of intra-observer variability for the manual calcium scoring results, in order to determine the correlation of manual and automatic scoring more precisely.

Overall, one of the major burdens when performing studies to the applicability of CTCA scoring for cardiac diagnostics, the increased manual effort, can be avoided by the use of our algorithm. Moreover, should those studies finally verify the clinical value of CTCA calcium scores, the presented method could become a valuable diagnostic tool for clinical practice.

References

1. Glodny B, Helmelt B, Trieb T, et al. A method for calcium quantification by means of CT coronary angiography using 64-multidetector CT: very high correlation with agatston and volume scores. *Eur Radiol.* 2009;19:1661–68.
2. Hoffmann MHK, Shi H, Schmitz BL, et al. Noninvasive coronary angiography with multislice computed tomography. *J Am Med Assoc.* 2005;293(20):2471–78.
3. Hadamitzky M, Freißmuth B, Meyer T, et al. Prognostic value of coronary computed tomographic angiography for prediction of cardiac events in patients with suspected coronary artery disease. *Int J Cardiovasc Imaging.* 2009;2(4):404–11.
4. Wesarg S, Khan MF, Firlie EA. Localizing calcifications in cardiac CT data sets using a new vessel segmentation approach. *J Digit Imaging.* 2006;19(3):249–57.
5. Gülsün MA, Tek H. Robust vessel tree modeling. *Proc MICCAI.* 2008; p. 602–11.
6. Agatston AS, Janowitz WR, Hildner FJ, et al. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol.* 1990;15:827–32.
7. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999;8:135–60.
8. Budoff MJ, Achenbach S, Blumenthal RS, et al. Assessment of coronary artery disease by cardiac computed tomography: a scientific statement from the american heart association committee on cardiovascular imaging and intervention, council on cardiovascular radiology and intervention, and committee on cardiac imaging, council on clinical cardiology. *Circulation.* 2006;114:1761–91.

Modellbasierte Echtzeit-Bewegungsschätzung in der Fluoreszenzendoskopie

Thomas Stehle, Jonas Wulff, Alexander Behrens, Sebastian Gross, Til Aach

Lehrstuhl für Bildverarbeitung, RWTH Aachen University, 52056 Aachen
thomas.stehle@lfb.rwth-aachen.de

Kurzfassung. In diesem Beitrag wird eine Methode zur echtzeitfähigen Block-Matching-basierten Bewegungsschätzung beschrieben, die auch in der Präsenz von starkem Rauschen zuverlässige Ergebnisse liefert. Zu diesem Zweck werden zunächst Blöcke kreisförmig im Bild verteilt. Mit Hilfe des Struktur-Tensors wird entschieden, ob ein Block genügend Strukturinformation aufweist, um zur Bewegungsschätzung herangezogen zu werden. Die ausgewählten Blöcke werden verwendet, um die Parameter eines globalen Bewegungsmodells unter Verwendung des RANSAC-Algorithmus robust zu bestimmen.

1 Einleitung

In Deutschland werden pro Jahr laut Robert Koch Institut etwa 28.750 neue Fälle von Blasenkrebs diagnostiziert. Diese Krebsart ist mit normaler Weißlicht-Endoskopie nur schwer zu festzustellen. Aus diesem Grund wird häufig die photodynamische Diagnose (PDD) angewandt, bei der dem Patienten eine Marker-substanz verabreicht wird, die sich vorwiegend an bösartigem Gewebe anlagert. Bei Verwendung einer schmalbandigen blauen Lichtquelle beginnen diese Areale zu fluoreszieren und der Tumor wird deutlich sichtbar. Ein Beispiel für Weißlicht- und PDD-Endoskopie wird in Abb. 1 gegeben. Die Tumore sind in beiden Fällen mit schwarzen Kreisen markiert. Unter Weißlicht sind die Tumore kaum zu erkennen, während sie sich bei der PDD deutlich von der gesunden Blasenwand unterscheiden.

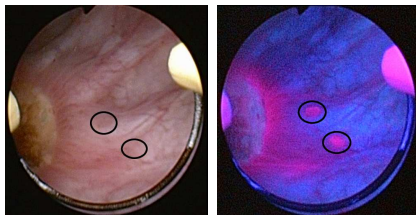


Abb. 1. Blasentumore unter Weißlicht- (links) und PDD-Endoskopie (rechts). Unter Weißlicht sind die Tumore (schwarze Markierung) schwer zu erkennen, während sie in der PDD deutlich sichtbar sind.

Durch die schmalbandige Beleuchtung und die lichtschwache Fluoreszenz trifft nur eine geringe Lichtmenge auf dem Videosensor des Endoskops auf. Aus

diesem Grund ist in solchen Videosequenzen meist ein starkes Rauschen präsent. Stehle et al. haben in [1] ein Verfahren vorgestellt, das eine echtzeitfähige Rauschreduktion in der PDD-Endoskopie ermöglicht. Dieses Verfahren setzt allerdings ein vorgegebenes Bewegungsvektorfeld voraus. Bewegungsschätzer, die zur Videokodierung eingesetzt werden, erzeugen nicht notwendigerweise ein glattes Bewegungsvektorfeld. Die Verwendung solcher Felder maximieren zwar die Prädiktionsleistung des Videocoders, bei der zeitlichen Filterung führen derartige Felder jedoch zu störenden Artefakten. Behrens hat in [2] ein Verfahren vorgestellt, das zur robusten modellbasierten Bewegungsschätzung zwischen zwei PDD-Bildern geeignet ist. Das Verfahren verwendet allerdings nur eine zeitlich unterabgetastete Folge der Eingangsbilder und ist trotzdem nicht echtzeitfähig.

Im Folgenden wird ein Verfahren zur modellbasierten Bewegungsschätzung vorgestellt. Das Verfahren ist geeignet, die benötigten Bewegungsvektoren in Echtzeit zu berechnen, so dass eine zeitlich rekursive Rauschfilterung in Echtzeit ermöglicht wird.

2 Material und Methoden

Das Videomaterial des Endoskopiesystems weist eine Auflösung von 768×576 Bildpunkten auf. Jedes Pixel besteht aus den drei Farbwerten rot, grün und blau und das System hat eine Aktualisierungsrate von 50 Hz. Das bedeutet, dass für eine Bewegungsschätzung (gefolgt von einer bewegungskompensierten zeitlich rekursiven Filterung) maximal 20 ms an Rechenzeit pro Bild zur Verfügung stehen, wenn eine echtzeitfähige Verarbeitung erwünscht ist.

Es wurde ein Verfahren entworfen, das ein globales Bewegungsmodell mit Hilfe von Punktkorrespondenzen bestimmt. Diese werden mit Hilfe eines Block-Matching-Verfahrens ermittelt. Die Verwendung eines solchen globalen Bewegungsmodells wird dadurch gerechtfertigt, dass sich das Endoskop meist sehr nah an der Blasenwand befindet, so dass diese in sehr guter Näherung als Ebene beschrieben werden kann.

2.1 Bestimmung der Punktkorrespondenzen

Eine Anzahl von Blöcken wird kreisförmig im Bild angeordnet und die entsprechenden Regionen werden im vorangegangenen Bild gesucht. Die Zuordnung der Mittelpunkte dieser Regionen sind die gesuchten Punktkorrespondenzen $[(x_i, y_i), (x'_i, y'_i)]$. Als Ähnlichkeitsmaß wird die Summe der absoluten Differenzen verwendet, da sich diese mit Hilfe der Befehlssätze heutiger Prozessoren sehr effizient berechnen lässt. Da Block-Matching mit Blockgrößen von 8×8 oder 16×16 Pixeln, wie sie in der Videokodierung angewandt werden, aufgrund des starken Rauschens kein robustes Ergebnis liefern, mussten in diesem Fall wesentlich größere Blöcke gewählt werden. Für diese Anwendung haben sich Blöcke der Größe 40×40 als stabil erwiesen. Auf der linken Seite von Abb. 2 ist das Block-Matching-Setup zu sehen.

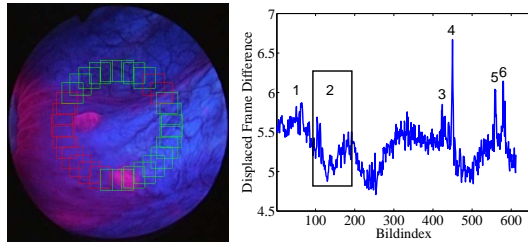


Abb. 2. Links: Anordnung der Such-Blöcke im Endoskopiebild. Die roten Blöcke wurden mit Hilfe des Strukturensors von der Bewegungsschätzung ausgeschlossen. Rechts: Darstellung der DFD für jedes Bild der Testsequenz. Signifikante Ereignisse wurden mit Ziffern gekennzeichnet.

Um die Zuverlässigkeit potentieller Korrespondenzen abzuschätzen, wird vor dem eigentlichen Block-Matching der Struktur-Tensor [3] jedes Blocks bestimmt. Die Eigenwerte des Struktur-Tensors geben ein Maß dafür an, wie viel Struktur in einem Block enthalten ist. Zwei kleine Eigenwerte deuten auf einen strukturlosen Block hin, der ggf. von Rauschen dominiert wird. Ein großer und ein kleiner Eigenwert lassen auf eine einfach gerichtete Struktur wie einen geraden Kantenzug schließen. Auch in diesem Fall ist aufgrund des Aperturproblems keine eindeutige Bewegungsschätzung möglich. Erst Blöcke mit zwei betragsmäßig großen Eigenwerten weisen auf Strukturen wie beispielsweise Ecken hin, die eine zuverlässige Bestimmung der Verschiebung erlauben.

2.2 Globale Bewegungsschätzung

Das globale Bewegungsmodell weist vier Freiheitsgrade auf: Translation in x - und y -Richtung (t_x und t_y), Rotation um den Winkel α und Skalierung um den Faktor s . Mit Hilfe dieser Freiheitsgrade kann ein Großteil der in der PDD-Endoskopie auftretenden Bildbewegungen modelliert werden. Durch die hohe zeitliche Abtastung ist der Einfluss anderer Bewegungen wie z. B. Scherung oder perspektivische Verzerrung zwischen zwei zeitlich benachbarten Bildern gering. Die entsprechende Abbildungsgleichung lautet

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} s \cdot \cos(\alpha) & -s \cdot \sin(\alpha) & t_x \\ s \cdot \sin(\alpha) & s \cdot \cos(\alpha) & t_y \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (1)$$

Mit Hilfe der Punktkorrespondenzen lässt sich nun folgendes lineares Gleichungssystem aufstellen, das eine Least-Squares-Lösung der Parameter des Bewegungsmodells liefert.

$$\begin{pmatrix} x'_0 \\ y'_0 \\ \vdots \\ x'_n \\ y'_n \end{pmatrix} = \begin{pmatrix} x_0 & y_0 & 1 & 0 \\ y_0 & -x_0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & 1 & 0 \\ y_n & -x_n & 0 & 1 \end{pmatrix} \begin{pmatrix} s \cdot \sin(\alpha) \\ s \cdot \cos(\alpha) \\ t_x \\ t_y \end{pmatrix} \quad (2)$$

Da sich trotz der Selektion der Blöcke noch Ausreißer in den Punktkorrespondenzen befinden können, wird der RANSAC-Algorithmus [4] zur robusteren Least-Squares-Schätzung der Modellparameter angewandt.

2.3 Experimente und Ergebnisse

Die Auswertung des vorgestellten Verfahrens wurde auf einer Endoskopiesequenz durchgeführt, bei der ein Arzt das Endoskop vom Ostium zur linken Blasenwand schwenkt. In dieser Sequenz sind neben Translation, Rotation und Skalierung aufgrund der Freihandbewegung ebenfalls perspektivische Bewegungskomponenten enthalten. Während der kompletten Sequenz sind bewegten sich kleine losgelöste Gewebepartikel durch das Bild, die nicht dem globalen Bewegungsmodell folgen und somit die Bewegungsschätzung erschweren.

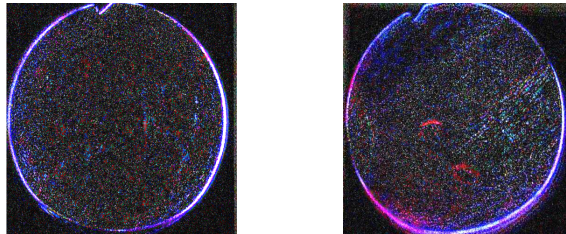
Als Bewertungsmetrik wurden die absoluten Differenzen zwischen einem zentralen Bereich des aktuellen Bilds und des entsprechend Gl. (1) bewegungskompensierten Bilds des vorherigen Zeitpunkts berechnet (displaced frame difference, DFD). Die DFD wird auf die Größe des zentralen Bereichs und auf die Anzahl der Farbkanäle normiert. Die Verwendung dieser Fehlerfunktion ist durch die zeitlich rekursive Filterung motiviert, für die das Verfahren entwickelt wurde. Ein niedriger Wert spricht für eine gute Bewegungskompensation und führt später zu einer starken Filterwirkung.

Der Verlauf der DFD während der Sequenz ist auf der rechten Seite von Abb. 2 zu sehen. Während der kompletten Sequenz findet eine kontinuierliche Bewegung statt. Verschiedene Ereignisse wurden dabei mit Ziffern gekennzeichnet. An der Stelle '1' wurde das Endoskop rotiert und eine Bewegungsunschärfe tritt auf. Das Verfahren erkennt die Bewegung korrekt und die DFD steigt nicht. Im Bereich '2' ist rechts im Bild eine Falte in der Blasenwand zu sehen, so dass die Hypothese einer einzigen globalen Bewegung verletzt ist. Die Inkonsistenz wird vom RANSAC korrekt erkannt und ausgeschlossen. Die DFD steigt hier nicht, da der Randbereich, in dem die Falte auftaucht, nicht zur DFD-Berechnung verwendet wird. Bei '3' findet eine ruckartige Rotation mit Bewegungsunschärfe statt, die das Verfahren nicht komplett erfasst. An den Stellen '4'-'6' treten sehr schnelle Zooms auf, die zu relativ großen Fehlern führen. Beispiele der DFD für eine erfolgreiche und eine ungenaue Bewegungsschätzung finden sich in Abb. 3. Das Verfahren benötigt auf einem 4 Kern Intel XEON Rechner mit 2,3 GHz ca. 18 ms zur Bewegungsschätzung, so dass für eine nachfolgende Filterung nur noch 2 ms zur Verfügung stehen. Da die Filterung nochmals etwa 12-15 ms benötigt, wird so keine Echtzeitfähigkeit erreicht. Durch die Anwendung eines Greedy-Optimierungsverfahrens [5] lässt sich die Rechenzeit auf 4 ms reduzieren, so dass eine Echtzeitverarbeitung ermöglicht wird. Die DFD nimmt dadurch im Mittel um etwa 1,2% zu.

3 Diskussion

Es wurde ein Verfahren zur echtzeitfähigen, modellbasierten Bewegungsschätzung in PDD-Endoskopie vorgestellt. Entsprechend der Anforderungen an das

Abb. 3. Links: DFD bei erfolgreicher Bewegungsschätzung. Bis auf die Randbereiche des Sichtfelds besteht die DFD vorwiegend aus Rauschen. Rechts: DFD bei ungenauer Bewegungsschätzung (Stelle '4' aus Abb. 2 rechts). Zusätzlich zum Bildrauschen sind Strukturinformationen sichtbar (2 Tumore und Blutgefäße der Blase, Abb. 2 links).



stark verrauschte Bildmaterial wurde ein Block-Matching-Verfahren angewandt, das aufgrund großer Blöcke robust gegenüber Rauschen ist. Die Blöcke wurden mit Hilfe des Struktur-Tensors auf Eignung zum Block-Matching hin untersucht und ggf. davon ausgeschlossen. Um die Robustheit gegenüber Ausreißern zu erhöhen, wurde der RANSAC-Algorithmus zur Bestimmung der Parameter des Bewegungsmodells verwendet.

Das Verfahren liefert gute Ergebnisse während langsamer Bewegungen und geringfügige Verletzungen der Annahme einer globalen Bewegung werden toleriert. Schnelle Zooms führen jedoch zu Fehlern in der Bewegungsschätzung.

Die Rechenzeit der vollständigen Suche konnte unter Verwendung eines Greedy Optimierungsverfahrens von 18 ms auf 4 ms reduziert werden, so dass noch ausreichend Rechenzeit für eine echtzeitfähige Filterung der Videosequenz zur Verfügung steht. Durch diese Maßnahme steigt die DFD im Mittel um 1,2% an und die Qualität der Bewegungsschätzung nimmt geringfügig ab.

Literaturverzeichnis

1. Stehle T, Wulff J, Behrens A, et al. Denoising fluorescence endoscopy: a motion-compensated temporal recursive video filter with an optimal minimum mean square error parametrization. In: Proc IEEE ISBI; 2009. p. 314–7.
2. Behrens A. Creating panoramic images for bladder fluorescence endoscopy. Acta Polytechn J Adv Eng. 2008;48(3):50–4.
3. Bigun J, Granlund GH. Optimal orientation detection of linear symmetry. In: Proc IEEE ICCV; 1987. p. 433–8.
4. Fischler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun ACM. 1981;24(6):381–95.
5. Williams DJ, Shah M. A fast algorithm for active contours and curvature estimation. Comput Vis Image Underst. 1992;55(1):14–26.

Planung und Simulation von Patchimplantaten zur intrakardialen Korrektur angeborener Herzfehler

Urte Rietdorf¹, Eugénie Riesenkampff², Tobias Schwarz¹, Titus Kuehne²,
Hans-Peter Meinzer¹, Ivo Wolf^{1,3}

¹Abteilung für Medizinische und Biologische Informatik, DKFZ Heidelberg

²Abteilung Kinderkardiologie und Angeborene Herzfehler, DHZB Berlin

³Institut für Medizinische Informatik, Hochschule Mannheim

`u.rietdorf@dkfz.de`

Kurzfassung. Chirurgische Korrekturen durch Rastelli-artige Operationen bei angeborenen Herzfehlern wie dem Double Outlet Right Ventricle (DORV) oder einer Transposition der großen Arterien (TGA) erfordern eine möglichst genaue Kenntnis der Anatomie des Herzens. Bei diesem Eingriff werden Fehlsprünge der großen Gefäße korrigiert, in dem mit Hilfe eines Patches eine tunnelartige Verbindung geschaffen wird, um so das Blut des linken Ventrikels in die Aorta zu leiten und gleichzeitig den Ventrikelseptumdefekt (VSD) zu verschließen. Die Planung des Patches muss unter Berücksichtigung der Anatomie erfolgen, um die Funktion nicht zu gefährden. So darf durch den Tunnel weder die Trikuspidalklappe noch der rechtsventrikuläre Ausflusstrakt behindert werden. Zur genaueren Visualisierung der Anatomie und zur Planung des Tunnels wurde daher eine Software entwickelt. Mit dieser ist es möglich, den Tunnel innerhalb des Herzens vom linken Ventrikel zum Gefäßsprung durch den VSD dreidimensional zu planen, interaktiv zu verschieben und sich abschließend einen zweidimensionalen Schnittplan zur Konstruktion des Tunnelpatches zu erstellen. Auf diese Weise können Prozeduren bezüglich ihrer technischen Machbarkeit präoperativ evaluiert werden.

1 Einleitung

Bei schweren Fehlbildungen des Herzens kann es durch eine fehlende Trennung des arteriellen und venösen Blutkreislaufes zu Funktionsbeeinträchtigungen der Herzfunktion kommen. Durch Anomalien des Ursprungs der großen Gefäße, wie sie beim Double Outlet Right Ventricle (DORV) oder einer Transposition der großen Arterien (TGA) vorkommen, sind die Blutflussverhältnisse oft so beeinträchtigt, dass ein normales Leben des Patienten nicht möglich ist. Zur chirurgischen Korrektur dieser Herzfehler gibt es die Option der Rastelli-artigen Operation. Hierbei wird der linke Ventrikel durch die Implantation eines Patchmaterials tunnelartig mit der Aorta durch den VSD verbunden (Abb. 1). Durch dieses Verfahren wird ein normaler Blutkreislauf etabliert.

Die Schwierigkeit dieser Operation liegt darin, die Lage und Form des einzusetzenden Patchmaterials so zu planen, dass wichtige Strukturen wie der Ausflusstrakt des rechten Ventrikels oder die Atrio-Ventrikuläre Klappe nicht beeinträchtigt werden. Derzeit erfolgt die Planung des Eingriffes einzig auf der Analyse der Bilddaten. Eine Softwarelösung zur Visualisierung der intrakardialen dreidimensionalen Struktur und zur Planung der Operation mit Patchvorschlag existiert nicht [1].

2 Material und Methoden

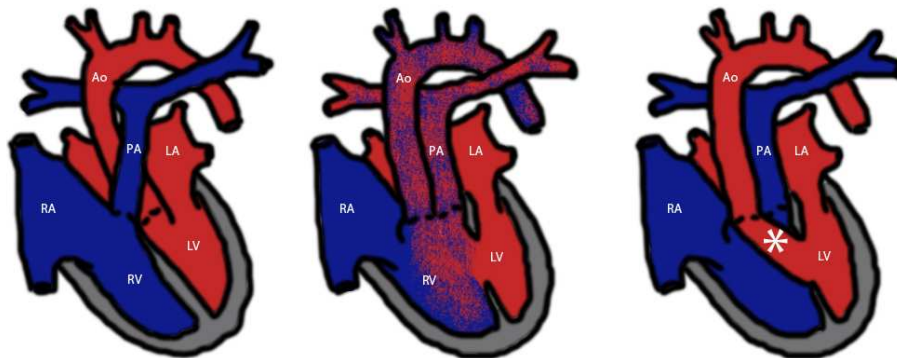
Um die Planung einer Rastelli-artigen Operation zu vereinfachen, wurde innerhalb des Medical Imaging und Interaction Toolkits MITK[2] eine Softwareapplikation erstellt. MITK ist ein Softwareframework zur Erstellung interaktiver medizinischer Applikationen, und basiert auf den Toolkits Imaging Toolkit ITK[3] und Visualization Toolkit VTK der Firma Kitware. Die Erstellung der Patchplanung erfolgt in drei Schritten:

1. Dreidimensionale Visualisierung der Herzstrukturen zur Analyse der Anatomie
2. Positionierung des Patchmaterials
3. Planung der Form des Patchmaterials

Abbildung 2 zeigt den Datensatz eines DORV vor der Planung eines Patches.

2.1 Visualisierung der Herzstrukturen

Die Orientierung innerhalb des schichtbasierten Herzdatensatzes zur Positionierung der Start- und Endbereiche des Patchmaterials kann sehr schwierig und



RA = Rechter Atrium; RV = Rechter Ventrikel; LA = Linker Atrium; LV = Linker Ventrikel; Ao = Aorta, PA = Pulmonalarterie, * = Tunnel

Abb. 1. Schematische Abbildung des Blutkreislaufs in einem gesunden Herzen (links), bei einem DORV (mitte), und nach einem Rastelli-artigen Eingriff (rechts).

aufwändig sein. Um die Positionierung des einzusetzenden Materials zu vereinfachen und somit eine adäquate Konstruktion zu ermöglichen, können die relevanten betroffenen Bereiche des Herzens, anhand derer geplant wird, als dreidimensionales Polygongitternetz dargestellt werden. Diese dreidimensionale Visualisierung ermöglicht das Setzen der Start- und Endpunkte des Patchmaterials in sowohl zwei- als auch dreidimensionalen Ansichten. Zur Generierung des Polygonmodells wird hierzu auf einem zuvor mittels eines Bounding Objects[4] definierten Ausschnittes des Datensatzes der schwellwertbasierte Marching-Cube Algorithmus[5] angewendet. Um zu vermeiden dass das Polygonnetz an den Bildkanten des definierten Bereiches abbricht, wird das neu erzeugte Bild der Region of Interest an den Rändern um einige dunkle Pixelreihen erweitert.

2.2 Positionierung des Tunnels

Soll ein Patchmaterial zur Konstruktion eines Tunnels eingesetzt werden, um so den Blutfluss des Herzens zu korrigieren, muss definiert werden, wo das Patchmaterial implantiert wird. Hierzu wird auf den erzeugten Oberflächenmodellen der Tunnelanfang und das Ende definiert, indem Punkte an den Start- und Endbereichen ringförmig um die Ausflussbereiche gesetzt werden. Dabei können die Punkte automatisch um einen Startpunkt herum gesetzt, anschließend jedoch manuell verschoben werden, um so neben der Position auch die Größe des zu erstellenden Tunnels zu beeinflussen. So kann beispielsweise unabhängig vom echten Ausmaß des VSDs bereits geplant werden, auf welchen Durchmesser er skaliert werden muss.

Zwischen Start- und Endring werden durch eine Pose-Interpolation[6] eine beliebige Anzahl weiterer ringförmig angeordneter Punkte erzeugt. Diese definieren die Tunnelwand. Nach Abschluss der Ringpositionierung werden alle den Tunnel definierenden Punkte in ein Oberflächenmodell integriert, welches dem Anwender den Tunnel dreidimensional visualisiert.

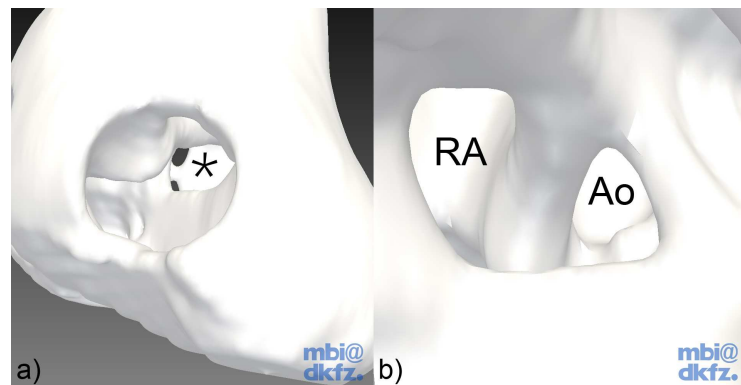


Abb. 2. Darstellung eines DORV. (a); Blick in den RV auf den VSD; (b): Blick aus dem LV durch den VSD auf die Ao.

2.3 Planung der Grösse und Form

Um den Patch entsprechend seiner Form und Größe vorbereiten zu können, muss ein zweidimensionaler Schnittplan des Tunnels erstellt werden. Hierzu werden entlang der Tunnelmittellinie die Umfänge ermittelt. Unter Berücksichtigung der Krümmung des Tunnels werden die Mittellinienpositionen und dazugehörigen Umfänge in ein Diagramm übertragen, und als Schnittvorschlag ausgegeben.

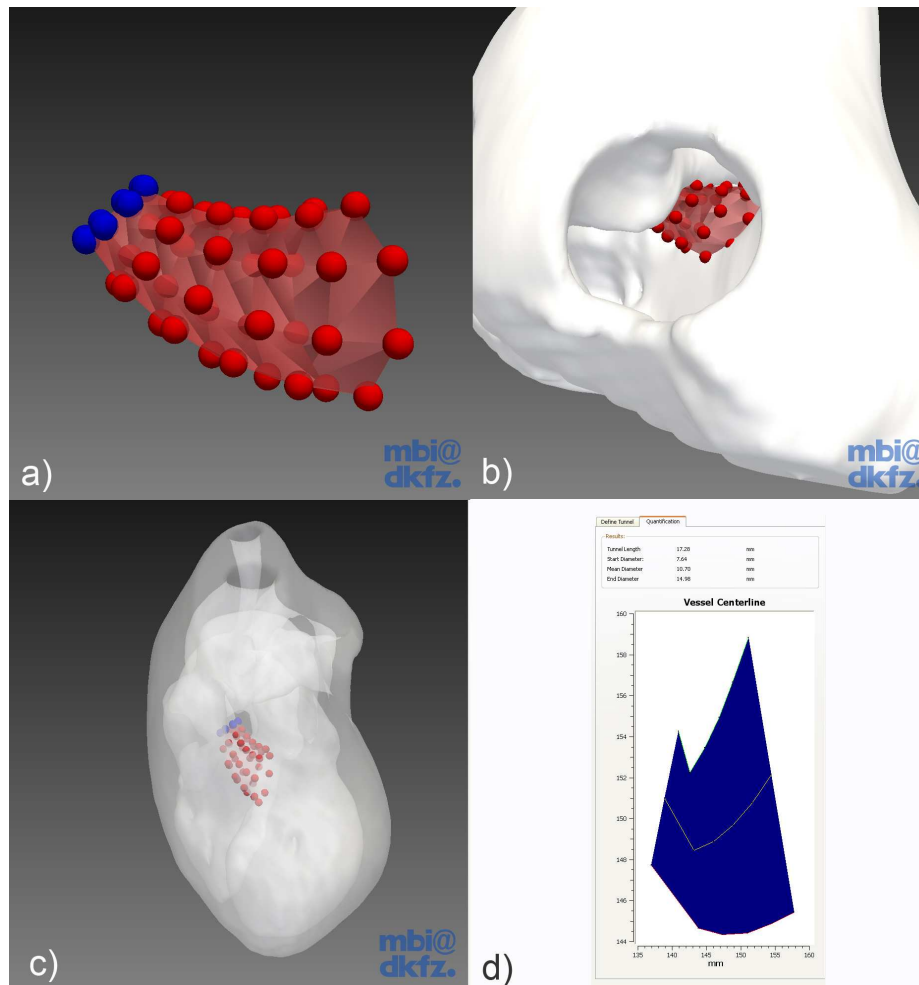


Abb. 3. Konstruierter Tunnel zwischen dem Aortenursprung und dem linken Ventrikel durch den Ventrikelseptumdefekt. (a): konstruierter Tunnel; (b): Ansicht von außerhalb des RVs; (c): Ansicht von schräg oben; (d): Quantifizierung des Tunnels und Patchvorschlag.

3 Ergebnis

Entstanden ist eine Software, mit der eine interaktive, dreidimensionale Planung eines Patches zur Konstruktion eines intrakardialen Tunnels ermöglicht wird. Dazu werden anhand der akquirierten MRT-Daten die vom Anwender definierte Bereiche des Herzen als dreidimensionale Oberflächenmodelle visualisiert. Anhand dieser kann anschließend der Verlauf des Tunnels definiert, manipuliert und visualisiert werden. Nach abgeschlossener Planung des Tunnels wird dieser in ein zweidimensionales Schnittmuster umgerechnet und ausgegeben, um eine Vorlage zur Erstellung des Patches zu liefern. Abbildung 3 zeigt den generierten Tunnel zwischen dem linken Ventrikel und der Aorta in verschiedenen Ansichten, sowie die Quantifikation des Tunnels und den zweidimensionalen Patchvorschlag.

4 Diskussion

Mit Hilfe der entwickelten Software ist es möglich, angeborene kardiale Anomalien wie DORV oder TGA als dreidimensionale Polygondaten darzustellen. Anhand dieser Visualisierung kann interaktiv ein Tunnel zur Rastelli-artigen chirurgischen Korrektur des Herzfehlers geplant und simuliert werden. Der entstandene Tunnel wird quantifiziert und als Patchvorschlag visualisiert, um den Chirurgen bereits präoperativ bei der Planung des Eingriffs zu unterstützen. Dadurch soll eine Prozedur bezüglich ihrer technischen Machbarkeit präoperativ evaluiert werden und idealer Weise die intraoperative Eingriffszeit vermindert und Risiken vermieden werden.

Literaturverzeichnis

1. Schumacher G, Hess J, Bühlmeier K. Klinische Kinderkardiologie Diagnostik und Therapie der angeborenen Herzfehler. 4th ed. Heidelberg: Springer Medizin Verlag; 2008.
2. Wolf I, Vetter M, Wegner I, et al. The medical imaging interaction toolkit (MITK). *Med Image Anal.* 2005;9:594–604.
3. Ibanez L, Schroeder W. The ITK Software Guide. Insight Software Consortium; 2003.
4. Vetter M, Neuhaus J, Wegner I, et al. Bounding-object segmentation. *Proc SPIE.* 2005;5747:1628–35.
5. Lorensen WE, Cline HE. Marching Cubes: A high resolution 3D surface construction algorithm. *Com Graph.* 1987;21(4):163–9.
6. Rossignac JR, Kim JJ. Computing and visualizing pose-interpolating 3D motions. *Computer-Aided Design.* 2001;33(4):279–91.

Multimodal Medical Consultation for Improved Patient Education

Patrick Wucherer¹, Christoph Bichlmeier¹, Maximilian Eder², Laszlo Kovacs²,
Nassir Navab¹

¹Chair for Computer Aided Medical Procedures & Augmented Reality,
Computer Science Department, TUM, Germany

²Department of Plastic and Reconstructive Surgery,
Klinikum rechts der Isar, TUM, Germany

p.wucherer@mytum.de

Abstract. Augmented reality (AR) being applied for preoperative patient education can become a major feature of new communication standards in medical consultation. We have analyzed the current workflow of patient education with respect to breast reconstruction in plastic surgery. According to the gathered information, we developed a concept for an AR supported patient education system and present first implemented components of this system. In addition, we summarize the qualitative feedback from interviews with surgeons.

1 Introduction

The majority of recent publications in the field of medical augmented reality (AR) preferably proposes this technology to intraoperatively support surgical procedures. Authors suggest this technology for example to reduce the focuses of attention e.g. in surgical navigation and to facilitate the mental mapping of imaging data onto the patient [1, 2]. We started an interdisciplinary collaboration between computer scientists and plastic surgeons in order to develop an AR system to support physicians to explain patients the limitations, risks and potentials of an intervention, the surgical procedure itself, potential complications, and prescription for preventive and postoperative behavior, treatment and medication. It is necessary for a patient to completely understand the treatment to consider and communicate the personal value they place on the benefits versus the harms [3].

Today's standard media providing the patient with such information is a printed brochure. Both, the physician and the patient review together the items of this brochure. An example of such brochures being used in Germany are the information and consent sheets of the company proCompliance (Thieme Compliance GmbH) or Perimed (perimed Fachbuch Verlag). In fact in many cases time is scheduled too shortly to ensure a complete, informative conversation and information is often too complex for the standard patient.

This work proposes a multi-modal communication tool for patient education that includes AR visualization.

2 Method

Our target system for patient education illustrated in Fig. 1(a, b) is a digital mirror having also been introduced as “magic mirrors” [4, 5] or virtual mirrors. Such mirrors usually consist of a projection screen and a video camera attached to the screen. The video data showing the environment in front of the screen is then projected onto the screen to create a mirror effect. Virtual objects can be registered with a person standing in front of the screen.

The system focuses on supporting the first and the second stage of the therapy workflow, when the patient has to understand complex information, decides on the procedure and gives the surgeon the approval of being aware of all legal information. We believe that the transfer of complex information and communication barriers caused by medical terminology being usually not familiar to the standard patient can be better managed when the briefing procedure currently performed with conventional spoken, written and 2D illustrated explication is augmented with information presented in-situ on the patient’s body.

Our expandable software framework provides features for interaction, multi-modal communication and virtual breast augmentation. In our preliminary version of registering the virtual breast and further virtual objects with patient’s anatomy, the patient wears an optical tracking target on her shoulder, which is located with the SmartTrack system of A.R.T. Weilheim GmbH. Markerless tracking of the patient’s upper body will be addressed in future stages of the project.

Since the system wants to augment but not replace traditional methods of patient education, conventional media presenting instructional information such as text and illustration can be displayed on the screen next to the AR scene.

The interaction of the system is based on a tracked remote mouse controller. Voice recognition (Microsoft Speech SDK) is used as an alternative interactive input option. The remote mouse controller also serves as a virtual 3D marker to distinguish specific regions on the patient’s skin.

The augmentation of the patient with a virtual breast model (Fig. 1(c)) currently involves three major features, which are the approximation of the biomechanical breast deformation, the composition of the virtual breast and the video



Fig. 1. Augmenting the patient with instructional information.

and the realistic texturing of the breast. The first step of the breast augmentation pipeline is the acquisition of a 3D data set of the patient's existing breast with a 3D scanner (Minolta VI-900). The scan is taken from a frontal position and another two profile positions by turning the patient to the right and to the left, 30 degrees each as proposed by Kovacs et al. [6]. Then the data is processed to generate a watertight breast volume for its tetrahedralization. The resulting model can be mirrored to obtain the breast to be reconstructed. We use Tetgen (<http://tetgen.berlios.de/>) to generate a volumetric mesh consisting of tetrahedral elements.

The mesh is animated using a GPU-computed mass-spring model [7]. For the integration step, we added artificial damping to enlarge the time step in order to guarantee the stability of the system. In addition, the early-z culling using a depth buffer is replaced by a texture lookup in the pixel shader during the force calculation to avoid processing unbound vertices.

The mass-spring system can be stimulated by the moving patient. The tracking target on the patient's shoulder provides 3D pose and acceleration data to deform the virtual breast. In a future version, the tetrahedral elements shall be parameterized using biomechanical parameters, e.g. proposed by Bianchi et al. [8]. The current version still uses dummy deformation properties for elements of the mass-spring system.

The breast model is colored with a skin texture generated with Adobe Photoshop. First an image of the whole upper body of the patient is loaded into the program. Then the existing breast is cut out of the image without the shaded regions. The mamilla is cut out and stored to a separate image layer. The mamilla image layer is moved by hand to its designated position in texture space. Finding this position currently bases on experience values. The border of the mamilla layer is blurred for a more smooth transition between skin texture and mamilla. The automatic generation of a breast texture including the mamilla is part of future work. The mapping of the texture onto the breast model bases on the projection from view uv unwrapping texture coordinates calculation done by Blender (blender.org). The uv-coordinates are drawn from the same point of view as the surface vertex coordinates of the 3D breast model.

For superimposing the virtual breast onto the patient's body we propose a processing pipeline for composition that exploits vertex buffer objects (VBOs), color buffer objects (CBOs) and frame buffer objects (FBOs).

The vertices, normals and colors of the surface of the breast model are stored to an index array defining a triangle structure. Each of these three items feeds one of three textures (VTex, NTex, CTex) of a FBO while the pixel index of each of the texture defines the correlation of the data in those textures.

In the first step, the normals of NTex belonging to the concave side of the breast shape are determined. Here, we assume that the model is positioned parallel to the xy-plane. For this reason, all normals on the surface of the 3D model are compared with a vector parallel to the z axis. Colors in CTex are then used to label corresponding normals in NTex and vertices in VTex by

indexes when normals have an acute angle with the mentioned vector (Step: Mark Vertices).

In the second step, the VTex and CTex of the FBO are used as VBO and as CBO respectively. Because of drawing only labeled vertices the resulting texture shows only the concave region of the breast surface.

After that, the edge of the concave surface is detected using a Sobel Filter and the border of the surface is extended by a dilation. Then the border is blurred to create a smooth blending mask between a texture sample taken of the video frame and the video frame itself.

A pointer based registration procedure finds the correspondence of four spots on the patient's skin and four positions in the volumetric mesh [9]. While the positions on the skin are selected during the consultation with a tracked remote mouse controller, the four positions in the volumetric mesh can be selected in a previous stage when the breast model is being prepared. The order of selecting the positions has to be the same for both resources.

3 Results

The feedback of five interviewed surgeons indicates that there is a demand for advanced communication tools for patient education. All surgeons mentioned the patient's major interest in the risks and consequences of the therapy. The combination of text, images of the patient herself and the virtual breast augmentation in one system would cover every aspect of patient education. Different communication strategies and levels of detail can be achieved to individualize the information to be transferred.

Surgeons favored the tracked mouse controller serving as a virtual pen to interactively augment the patient with virtual 3D sketches. The virtual sketches can be used to explain the procedure and its consequences, for instance scars that may occur, but also to extend the documentation of the clinical workflow and to measure anatomic distances.

Most of the surgeons would like to use this system to present the outcome of the surgery, provided that the prediction is realistic. This also involves breast reconstruction as well as breast augmentation in aesthetic surgery. However, the realistic prediction strongly relates to the exact simulation of the biomechanical parameters of the breast tissue, which is still subject of ongoing research. However, surgeons mention that they would already use the system for patients having small breasts with less deformation.

Along with the desire to provide only information that is honest and ethically correct, some surgeons have recommended to display the anatomy rather with an abstract than realistic visualization.

4 Discussion

We presented the concept and first setup of a novel communication tool that augments the current media standard for patient education. According to feedback

from interviewed medical partners the system can be useful to explain the limitations, risks, the surgical procedure itself, potential complications and postoperative scars, preventive information and prescription for postoperative behavior, treatment and medicamentation. We plan to replace the optical tracking system with a TOF (Time of flight) camera to enable interaction as well as anatomy registration and increase the usability of the system.

In addition, biomechanical information for realistic deformation has to be integrated once research has progressed in this area. For more realistic simulation of soft tissue, an advanced deformation model has to be applied that that for example may be based on the Finite Element Method proposed by Bianchi et al. [8]. Future work also includes an advanced method of composing the virtual breast and the patient. We currently plan to do early patient studies in order to better understand further requirements of the system.

Acknowledgement. We would like to thank A.R.T. GmbH for providing tracking cameras and software. We also want to express our gratitude to Christian Brossmann and to the crew at NARVIS lab.

References

1. Bianchi, G, Solenthaler, et al. Simultaneous topology and stiffness identification for mass-spring models based on FEM reference deformations. In: Proc MICCAI; 2004. p. 293–301.
2. Birkfellner W, Figl M, Matula C, et al. Computer-enhanced stereoscopic vision in a head-mounted operating binocular. *Phys Med Biol.* 2003;48(3):49–57.
3. Echtler F. Realization of efficient mass-spring simulations on graphics hardware [PhD Thesis]. Technische Universität München, Department of Computer Science; 2004.
4. Fiala M. Magic mirror system with hand-held and wearable augmentations. In: VR. IEEE Computer Society; 2007. p. 251–4. Available from: <http://dblp.uni-trier.de/db/conf/vr/vr2007.html#Fiala07>.
5. Kovacs L, Yassouridis A, Zimmermann A, et al. Optimization of 3-Dimensional Imaging of the Breast Region With 3-Dimensional Laser Scanners. Department of Plastic and Reconstructive Surgery and Department of Surgery, Klinikum rechts der Isar of the Technical University Munich; 2006.
6. Michalke N. Trying on clothes in a magic mirror. *Fraunhofer Magazin.* 2009;1:28.
7. O'Connor AM, Rostom A, Fiset V, et al. Decision aids for patients facing health treatment or screening decisions systematic review. *Br Med J.* 1999;319:731–4.
8. Walker MW, Lejun S, Volz RA. Estimating 3-D location parameters using dual number quaternions. *CVGIP: Image Underst.* 1991;54(3):358–67.
9. Wacker FK, Vogt S, Khamene A, et al. An augmented reality system for MR image-guided needle biopsy: Initial results in a swine model. *Radiology.* 2006;238(2):497–504.

Praktische Aspekte zur hochauflösenden Stereovideo-Dokumentation intraoperativer Befunde in der Hals-, Nasen- und Ohrenheilkunde

Justus Ilgner, Jonas Jae-Hyun Park, Martin Westhofen

Klinik für Hals-, Nasen- und Ohrenheilkunde und Plastische Kopf- und Halschirurgie
der RWTH Aachen
jilgner@ukaachen.de

Kurzfassung. Die stereovideografische Dokumentation intraoperativer Befunde ist mit der Digitalisierung der Bildverarbeitungsprozesse ein viel versprechender Ansatz zur Verbesserung der medizinischen Ausbildung. Wir verwenden ein mikroskopisches Kamerasystem mit 1280×720 Pixel Auflösung, welches das Signal beider Kanäle über eine PC-Workstation mit einem RAID-0-Array bei 18 fps aufzeichnet. Die Prozesszeiten im Op werden durch den Einsatz des Systems initial um durchschnittlich 12 bis 17 Minuten verlängert. Jedoch wird die Ausbildung junger Operateure wesentlich dadurch erleichtert, dass eine kontinuierliche Überwachung der Op durch den Ausbilder stattfindet und häufige Wechsel zwischen Ausbilder und Auszubildendem vermieden werden.

1 Einleitung

Seit der Einführung des Operationsmikroskops in den Operationssälen Ende der 1950er Jahre und der Entwicklung lichtstarker starrer und flexibler Endoskope seit ca. 40 Jahren ist der relative Anteil der Operationen ohne optische Hilfsmittel stetig zurückgegangen. Der allgemeine Wunsch nach minimal-invasiven Operationstechniken verstärkt diesen Trend kontinuierlich. So wurden in unserer Klinik von 2048 Operationen eines Jahres 45 % aller Operationen ohne optische Hilfsmittel durchgeführt, während sich der Anteil endoskopischer Eingriffe auf 6,7 % und mikroskopischer Operationen auf die übrigen 48,3 % belief.

Die Vermittlung und Kontrolle minimal-invasiver Op-Techniken an jüngere Kollegen stellt für erfahrene Operateure eine bedeutende Herausforderung dar [1, 2]. Traditionell findet die Ausbildung mikroskopischer Eingriffe am optisch ausgekoppelten Bild statt, welche wegen der räumlichen Anordnung des Mikroskops nur ein monoskopisches Bild über den Seitenausgang eines Strahlenteilers liefert. Die Kontrolle über das Operationsgebiet ist nur dem Operateur möglich, so dass in der Ausbildung häufig die Positionen am Patienten gewechselt werden müssen. Der Nutzen stereoskopischen Videomonitorings ist in vorangegangenen Arbeiten anhand laparoskopischer Operationen untersucht worden [3, 4]. Dennoch haben sich diese Monitortechniken am Markt nicht allgemein durchgesetzt.

Gegenstand der Untersuchung sind der Nutzen einer hochauflösenden stereovideoskopischen Monitoranlage im Operationssaal sowie die Randbedingungen für den praktischen Einsatz.

2 Material und Methoden

Das verwendete Kamerasystem besteht aus einem stereoskopischen Kamerakopf mit zwei parallel angeordneten 1CCD-Bildaufnehmern mit einer Auflösung von je 1280×1024 Pixel. Der Anschluss an das Mikroskop (OpMi proMagis, Zeiss, Oberkochen, Deutschland) erfolgt über eine Standard-Klemmvorrichtung entweder anstelle des Binokulartubus oder über einen stereoskopischen Strahlenteiler im Strahlengang unmittelbar hinter dem Mikroskopkopf. Die digitalen Bildsignale werden in einer PC-Workstation (zwei Intel® Xeon® CPU, 3.0 GHz, 1024 MB RAM; Intel Co., Santa Clara, CA, USA) verarbeitet. Die verwendete Software erlaubt Videoaufzeichnungen über ein RAID 0 Festplattenarray von 4×750 GB mit einer Gesamtaufzeichnungsdauer von 3 h 18 min bei 18 fps. Die Wiedergabe erfolgt online wahlweise über zwei LCD-Monitore mit 1280×1024 Pixel Auflösung, die in einem 120 Grad Winkel zueinander angeordnet eine simultane Betrachtung über einen an der Winkelhalbierenden angebrachten halbdurchlässigen Spiegel erlauben. Alternativ wird ein Rückprojektionssystem mit zwei Standard-LCD-Projektoren mit 1280×720 Pixel (Sharp Co., Tokio, Japan) verwendet. Für die Untersuchung wurden insgesamt 12 Operationen von 5 Operateuren ausgewertet, davon 4 mikroskopisch-endolaryngeale Eingriffe am Kehlkopf, 7 mikroskopisch-rekonstruktive Eingriffe am Mittelohr und ein mikroskopisch-endonasaler Eingriff.

Ausgewertet wurden die benötigte Zeit zum Auf- und Abbau des Systems, der subjektive Bildeindruck des Operateurs, des medizinischen Assistenzpersonals und der zusätzlichen Betrachter sowie die Auswirkungen auf den Arbeitsablauf.

3 Ergebnisse

Auswirkungen auf die Prozesszeiten: Der Aufbau des Systems im Operationssaal nahm für die Variante des LCD-Bildschirmarrays 12 Minuten (± 3 Minuten) sowie für die Rückprojektions-Variante 17 Minuten (± 5 Minuten) in Anspruch. Der Abbau des Systems geschah hingegen in beiden Varianten mit 7 bzw. 8 Minuten (± 2 Minuten) gleich schnell.

Subjektiver Bildeindruck der Beteiligten: Der allgemeine subjektive Bildeindruck wurde als „gut“ bis „sehr gut“ von allen 5 Operateuren eingestuft. Der stereovideografische Bildeindruck stellte sich innerhalb von weniger als 10 Sekunden bei allen Operateuren und bei 5 von 7 assistierenden Pflegekräften ein, bei einer Pflegekraft innerhalb von 30 Sekunden und bei einer weiteren (bekannter Astigmatismus) gar nicht. Eine Pflegekraft und ein weiterer Betrachter (1 Medizinstudent von 10) beklagten visuelle Ermüdung nach 10 bzw. 12 Minuten.

Prozessablauf: Im Fall einer mikroskopisch-rekonstruktiven Operation des Mittelohrs war keiner von 3 Ohroperateuren bereit, zugunsten des Videomonitorings auf eine direkte Sicht durch den Binokulartubus zu verzichten. Als Grund hierfür wurden besonders hohe Anforderungen an das räumliche Auflösungsvermögen sowie Kontrast und Farbwiedergabe angegeben. Hingegen waren drei der 4 mikroskopisch-endolaryngealen Eingriffe von 3 Operateuren unter ausschließlicher stereovideografischer Sicht durchgeführt worden (Abb. 1). Hierfür wurde die alleinige Wiedergabe über das Stereovideomonitoring als ausreichend angesehen, da der korrekte räumliche Eindruck keinen Nachteil gegenüber dem direkten binokularmikroskopischen Bild darstellte.

Im Fall der mikrolaryngoskopischen Eingriffe wurde die Positionierung von Assistenz, Operateur und Ausbilder in eine Linie vor der Projektionsfläche geändert (Vorher: Vis-à-vis) (Abb. 1, 2). Hierdurch reduzierte sich die Arbeitszeit eines 35minütigen Eingriffs (-11/+13 Minuten) um durchschnittlich 3 (± 1) Minuten, da der Arbeitsablauf vom Assistenzpersonal mitverfolgt und der nächste Schritt vorausgesehen werden konnte.



Abb. 1. Praktische Durchführung eines mikroskopisch-endolaryngealen Eingriffs unter Stereovideomonitoring als alleinige Bildwiedergabemodalität: Stereovideo-Kamerakopf (Pfeil), Operateur (links), Rückprojektionsfläche (Stern), Op-Assistenz (Mitte), Ausbilder (rechts).

4 Diskussion

Stereovideoskopische Monitoreinrichtungen im Operationssaal sind seit mehreren Jahren verfügbar [5], haben sich aber wegen des erhöhten Aufwandes beim Betrieb und Auf- und Abbau nicht allgemein durchgesetzt. Eine besondere Schwierigkeit besteht in der Synchronisation beider Videokanäle bei der analogen Aufzeichnung, wohingegen entsprechende Software für die synchrone stereoskopische Videoverarbeitung verfügbar ist [6]. Im hier vorgestellten System wird die zeitliche und räumliche Wiedergabequalität erhalten. Nachteilig ist die durch die Verwendung zweier Wiedergabegeräte erforderliche Justage der Monitore aufeinander zu Beginn jedes Aufbaus, der sich in den gegenwärtigen Prozesszeiten negativ niederschlägt. Bei kurzen Operationszeiten von bis zu ca. 30 Minuten (Panendoskopie mit Probenentnahme) wird der erhöhte Zeitaufwand zum Auf- und Abbau sowie Justage der Anlage nicht aufgewogen. Der erhöhte Zeitauf-

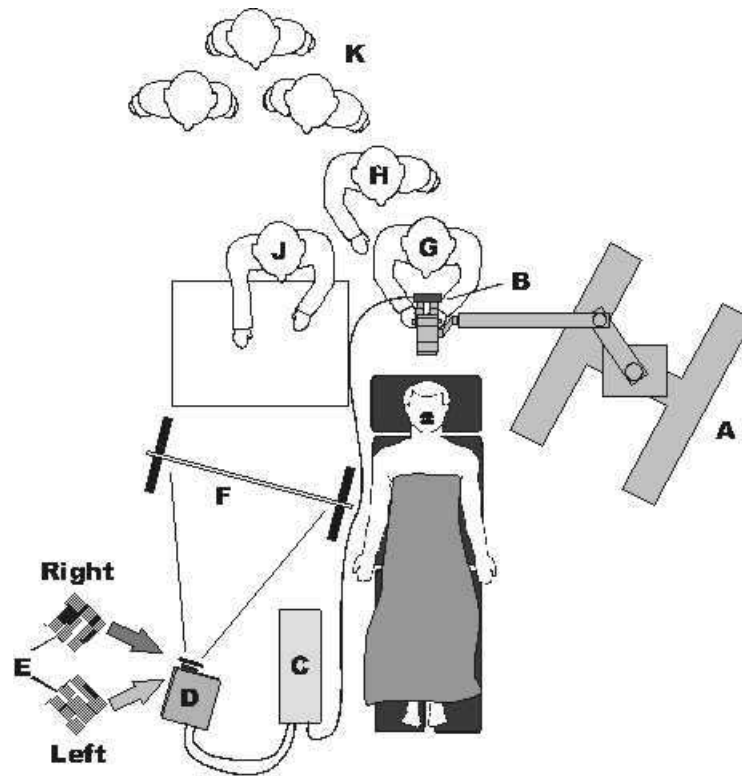


Abb. 2. Allgemeiner Aufbau der Stereovideoanlage im Operationssaal bei einem mikroskopisch-endolaryngealen Eingriff: a) Mikroskop, b) Stereo-Kamerakopf, c) PC-Workstation, d) 2 LCD-Projektoren, in Metallrahmen verbunden, e) Gegenläufige Orientierung der linearen Polfilter vor der Projektionslinse, f) Rückprojektionsfläche, g) Operateur, h) Ausbilder, j) Pflegepersonal, k) weitere Betrachter.

wand wird jedoch dann kompensiert, wenn junge Operateure ausgebildet werden sollen, da der häufige Wechsel von Ausbilder und Auszubildendem am Binokulartubus durch die für alle stereoskopisch verfügbare Kontrolle des Eingriffs entfällt. Dies reduziert das Patientenrisiko, da alle Beteiligten zu jedem Zeitpunkt über die gleiche dreidimensionale Bildinformation verfügen. Eine Steigerung der Präzision videostereoskopischer Eingriffe ergibt sich unmittelbar nicht, da diese von anderen Faktoren (Anatomie, manuelle Geschicklichkeit) wesentlich mitbestimmt wird, ist aber durch evtl. Kombination mit Navigationstechnologien denkbar. Als nächster sinnvoller Schritt aus Anwendersicht ergibt sich daher neben der gesteigerten Bildauflösung die selektive Verstärkung von Kontrast und Modifikation der Farbwiedergabe sowie die Integration CT-gesteuerter Navigationsanwendungen, um funktionell wichtige Strukturen frühzeitig erkennen und erhalten zu können [7].

Literaturverzeichnis

1. Ilgner J, Kawai T, Westhofen M, et al. Production and evaluation of stereoscopic video presentation in surgical training. *Proc SPIE*. 2004;5291:293–302.
2. Ilgner J, Kawai T, Shibata T, et al. Evaluation of stereoscopic medical video content on an autostereoscopic display for undergraduate medical education. *Proc SPIE*. 2006;6055:46–56.
3. von Pichler C, Radermacher K, Boeckmann W, et al. Three-dimensional versus two-dimensional video endoscopy. A clinical field study in laparoscopic application. *Stud Health Technol Inform*. 1996;29:667–74.
4. von Pichler C, Radermacher K, Boeckmann B, et al. The influence of LCD shutter glasses on spatial perception in stereoscopic visualization. *Stud Health Technol Inform*. 1996;29:523–31.
5. Meneses MS, Cruz AV, Castro IA, et al. Stereoscopic Neuroanatomy: comparative study between anaglyphic and light polarization techniques. *Arq Neuropsiquiatr*. 2002;60(3-B):769–74.
6. Kawai T, Shibata T, Inoue T, et al. Development of software for editing stereoscopic 3-D movies. *Proc SPIE*. 2002;4660:58–65.
7. Goes VP, Machado LS, Cabral MC, et al. Interactive stereoscopic full-color direct volume visualization for virtual reality applications in medicine. *Stud Health Technol Inform*. 2001;81:161–7.

Tracking von Gesichtsmimik mit Hilfe von Gitterstrukturen zur Klassifikation von schmerzrelevanten Action Units

Christine Barthold¹, Anton Papst¹, Thomas Wittenberg¹
Christian Küblbeck¹, Stefan Lautenbacher², Ute Schmid², Sven Friedl^{1,3}

¹Fraunhofer-Institut für Integrierte Schaltungen IIS, Erlangen,
²Otto-Friedrich-Universität Bamberg, ³Universitätsklinikum Erlangen
`sven.friedl@iis.fraunhofer.de`

Kurzfassung. In der Schmerzforschung werden schmerzrelevante Mimikbewegungen von Probanden mittels des Facial Action Coding System klassifiziert. Die manuelle Klassifikation hierbei ist aufwändig und eine automatische (Vor-)klassifikation könnte den diagnostischen Wert dieser Analysen erhöhen sowie den klinischen Workflow unterstützen. Der hier vorgestellte regelbasierte Ansatz ermöglicht eine automatische Klassifikation ohne große Trainingsmengen vorklassifizierter Daten. Das Verfahren erkennt und verfolgt Mimikbewegungen, unterstützt durch ein Gitter, und ordnet diese Bewegungen bestimmten Gesichtsarealen zu. Mit diesem Wissen kann aus den Bewegungen auf die zugehörigen Action Units geschlossen werden.

1 Einleitung

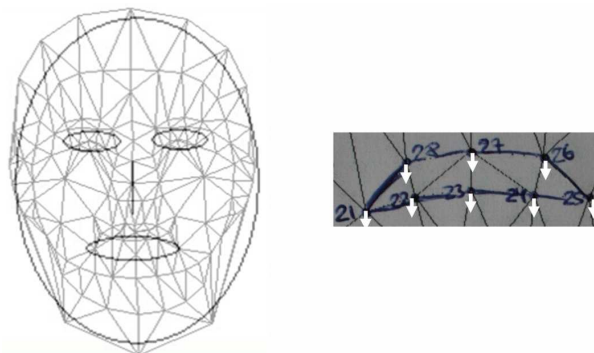
Menschliche Empfindungen wie Emotionen oder Schmerz lösen spezifische Muster von Kontraktionen der Gesichtsmuskulatur aus, die Grundlage dessen sind, was wir Mimik nennen. Aus der Beobachtung der Mimik kann wiederum auf menschliche Empfindungen rückgeschlossen werden. Im Rahmen der Schmerzforschung werden Videoaufnahmen von Probanden hinsichtlich des mimischen Schmerzausdrucks analysiert. Zur Beschreibung des mimischen Ausdrucks und dessen Veränderungen wird das Facial Action Coding System (FACS) [1] verwendet, das anatomisch begründet, kleinste sichtbare Muskelbewegungen im Gesicht beschreibt und als einzelne Action Units (AUs) kategorisiert. Eine Vielzahl von Untersuchungen hat gezeigt, dass spezifische Muster von Action Units auftreten, wenn Probanden Schmerzen angeben [2]. Die manuelle Klassifikation und Markierung der Action Units von Probanden in Videosequenzen bedarf einer langwierigen Beobachtung durch ausgebildete FACS-Coder. Eine automatische (Vor-)klassifikation kann hierbei den klinischen Workflow unterstützen und dieses Verfahren zum brauchbaren diagnostischen Instrument machen. Bisher realisierte Ansätze zum Erkennen von Gesichtsausdrücken basieren auf der Klassifikation

von trainierten Merkmalen, wie zum Beispiel Deformationsparametern [3] oder Grauwertverteilungen [4]. Neben der Notwendigkeit von großen Trainingsmengen für jede mögliche Ausprägung einer Action Unit, wird dem FACS-Coder auch der Zusammenhang zwischen der Klassifikation und der auslösenden Mimikänderung nicht deutlich erkennbar gemacht. Um Ausprägungen von Action Units ohne große Trainingsmengen klassifizieren zu können und eine nachvollziehbare Erklärungskomponente zu liefern, wird hier ein regelbasiertes Verfahren vorgestellt, das sich an den Beschreibungen des FACS orientiert.

2 Material und Methoden

Das FACS beschreibt, wie sich anatomische Positionsmerkmale im Gesicht, wie zum Beispiel Augenbrauen oder Mund, verändern müssen, damit eine bestimmte Action Unit gegeben ist. Zum Erlernen der Gesichtsausdrücke, die in Action Units kodiert werden, wird ein Regelsystem aufgebaut, welches den Beschreibungen der Positionsänderungen entspricht. Um die geforderten Positionsänderungen einer Action Unit zu erkennen, werden mimische Einzelbewegungen in Videosequenzen erkannt und über die Zeit verfolgt. Mit Hilfe einer Gitterstruktur, welche an das Gesicht angepasst wird, werden Bewegungen bestimmten Arealen im Gesicht zugeordnet (1). Die Gitterstruktur kann hierbei automatisch oder manuell an das Gesicht angepasst werden. Um Fehler aus der automatischen Anpassung zu verringern wurde in dieser Arbeit das Gitter anhand von 18 manuell gesetzten Landmarken geschätzt und angepasst. Insgesamt besteht das verwendete Gitter aus 151 Punkten. Treten mimische Einzelbewegungen auf, werden diese mittels Verfahren der Bewegungserkennung verfolgt. Ansätze hierzu sind Verfahren der Differentialbestimmung in Bildsequenzen wie der Optische Fluss [5] oder Verfolgung von Merkmalspunkten [6] über die Bildfolge. Die erkannten Bewegungen werden den anliegenden Gitterpunkten zugeordnet und das Gitter entsprechend neu angepasst. Da jedem der Gitterpunkte ein spezifisches Areal im Gesicht zugeordnet ist, kann aus der Gitterdeformation auf die aktivierte Muskelgruppe und dadurch auf die zugehörige Action Unit geschlossen werden.

Abb. 1. Skizze der an ein Gesicht angepassten Gitterstruktur (links) und Ausschnitt der Gitterstruktur an der rechten Augenbraue (rechts). Die Pfeile beschreiben die Bewegungsrichtung und -länge der Erscheinungsänderung *Augenbraue bewegt sich nach unten* (AU 4).



Beim Erstellen des Regelwerks werden Bewegungen an Gesichtsmerkmalen über die zugeordneten Gitterpunkte analysiert.

Das Regelwerk prüft für jede Erscheinungsänderung einer Action Unit folgende drei Aspekte. In Abb. 2 werden die ersten zwei Aspekte graphisch verdeutlicht.

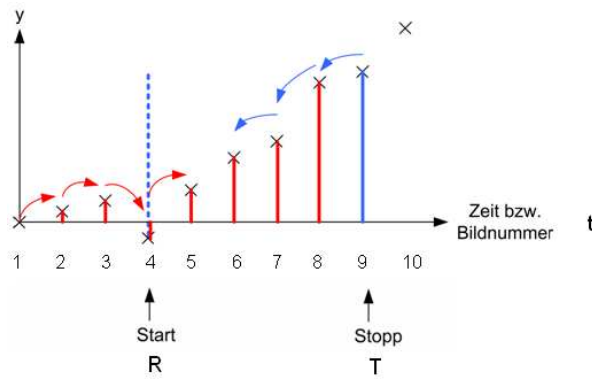


Abb. 2. Abfragepunkte des Regelwerks. Stopp- und Startmarke werden gesetzt, wenn die Bewegungslänge zwischen der Gitterposition im ersten und neunten Bild und zwischen neuntem und viertem Bild ausreichend groß ist.

1. *Bewegungsrichtung*: Die für die Entdeckung einer schmerzrelevanten mimischen Einzelbewegung betrachteten Gitterpunkte müssen sich in die durch die Action Unit beschriebene Richtung bewegen.
2. *Bewegungslänge*: Die Distanz der Bewegung des Gitterpunkts von der Ausgangsposition im ersten Bild der betrachteten Bildfolge zur Endposition in einem zeitlich späteren Bild der Bildfolge muss größer als ein entsprechend festgelegter Schwellwert θ_1 sein

$$\sum_{t=t_0}^T d_t > \theta_1$$

mit T : Bildnummer der Stopppmarke und d_t : Bewegungslänge eines Gitterpunkts zwischen Bild t und $t + 1$. Ist dies der Fall wird eine Stopppmarke T gesetzt.

3. *Bewegungsdauer*: Durch Rückwärtsprüfen der Bewegung von der Stopppmarke T bis zu einem zeitlich früheren Bild wird eine Startmarke R gesetzt. R wird zu dem Zeitpunkt oder Bild gesetzt, bei dem die Bewegungsdistanz im Intervall $I = T - R$ einen Schwellwert θ_2 überschreitet

$$\sum_{t=T}^R d_t > \theta_2$$

mit R : Bildnummer der Startmarke und $R \leq T$. Das Intervall I , das durch die Start- und Stopppmarke definiert wird, beschreibt die Dauer der betrachteten Bewegung.

3 Ergebnisse

Für diese Arbeit wurden exemplarisch vier schmerzrelevante Action Units implementiert. Dies sind die Action Units 4 (Brow Lowerer), 7 (Lid Tightener), 10 (Upper Lip Raiser) und 43 (Eyes Closed). Bisher werden nur einzeln auftretende Action Units und keine Kombinationen klassifiziert. Zur Evaluierung des Verfahrens wurde annotiertes Bildmaterial des Facial Action Coding System Manuals [1] und der Bilddatenbank von Kanade, Cohn und Tian [7] sowie eigens erstellte Bildsequenzen verwendet. Es hat sich gezeigt, dass die anatomischen Unterschiede zwischen einzelnen Personen die jeweiligen Bewegungen sehr unterschiedlich ausfallen lassen. Daher wurden die Schwellwerte der Bewegungsintensitäten für die verschiedenen Erscheinungsänderungen personenabhängig bestimmt. Da nur für die eigens erstellten Aufnahmen Informationen über die Bewegungsdauer vorhanden waren, wurde dieser Aspekt vernachlässigt. Der Schwellwert hierzu wurde so gewählt, dass das Intervall zwischen einem Bild und der maximalen Anzahl der vorhandenen Bilder der untersuchten Bildsequenz liegt. Mit dem vorhandenen Bildmaterial konnten in Experimenten 83 der 117 Testfälle erfolgreich klassifiziert werden. In den anderen Fällen wurde keine AU erkannt. Dies entspricht einer Klassifikationsrate von 70 %. Abbildung 3 zeigt ein Beispiel der Bewegungsverfolgung der Action Unit 4, welche erfolgreich klassifiziert werden konnte. Für die Berechnung des optischen Flusses wurde vorwiegend das Verfahren nach Horn und Schunck [5] verwendet. Hier konnten bei der experimentellen Untersuchung der Genauigkeit bessere Ergebnisse gegenüber dem Kanade-Lucas-Tomasi Feature Tracker [4] für Sequenzen mit Mimikbewegungen erzielt werden. Die wesentliche Fehlerquelle für eine nicht erfolgreiche Klassifikation lag in ungenauen Ergebnissen der Bewegungsanalyse. Die öffentlich verfügbaren Bilddaten [1, 7] sind nur bedingt für eine automatische Klassifikation mit diesem Ansatz geeignet. Eine sehr geringe Auflösung und verrauschte und teils unscharfe Bilder



Abb. 3. Eine Bildfolge aus [7] mit Mimikbewegungen der Action Unit 4. Erstes und letztes Bild der Action Unit sowie die erkannte Bewegung die zu einer erfolgreichen Klassifikation führte.

verhindern eine exakte Bewegungsverfolgung mit einer akzeptablen Genauigkeit. Bei Bewegungen von teilweise nur 5–10 Pixel im Bild führen bereits kleine rauschbedingte Fehler zu Distanzabweichungen von 50 % in der Bewegungserkennung. Konnte jedoch die Bewegung in den Bildsequenzen nicht genau genug erkannt und verfolgt werden, scheitert entsprechend auch das Regelsystem zur Klassifizierung. Mit dem hochwertigeren Bildmaterial konnten bessere Ergebnisse erzielt werden, jedoch standen nur wenige Aufnahmen zur Verfügung.

4 Diskussion

Es wurde ein regelbasierter Ansatz vorgestellt, der es ermöglicht Action Units des FACS zu klassifizieren ohne auf große Trainingsmengen vorklassifizierter Daten angewiesen zu sein. Zudem erlaubt dieser Ansatz den Zusammenhang zwischen kodierter Action Unit und zugrundeliegenden Änderungen von Positionsmerkmalen im Gesicht darzustellen. Beeinflusst wird das Ergebnis der Klassifikation von der Präzision der Anpassung der darunterliegenden Gitterstruktur sowie der Genauigkeit der Bewegungsverfolgung der Mimikbewegungen. Schwierigkeiten hierbei bereitet Bildmaterial, in dem das Gesicht nur sehr klein und unscharf abgebildet ist. In fortführenden Entwicklungen sollte das Regelsystem im Hinblick auf weitere Action Units und deren Erscheinungsänderungen erweitert werden. Zudem ist das gemeinsame Auftreten von mehreren Units die Regel und die Erkennung solcher Action Unit-Pattern diagnostisch besonders interessant, so dass auch hier ein weiterer Fokus gesetzt werden sollte. Eine Erweiterung des Regelsystems mittels eines Kantendetektors würde das Erkennen von Faltenbildung ermöglichen, was als Ergänzung im Regelwerk zu einer Verbesserung der Klassifizierung führen kann.

Literaturverzeichnis

1. Ekman P, Friesen WV, Hager JC. Facial Action Coding System The Manual [On CD ROM]. Nexus division of Network Information Research Corporation; 2002.
2. Kunz M, Mylius V, Schepelmann K, et al. On the relationship between verbal report and facial expression of pain. *J Pain*. 2004;5:368–76.
3. Black MJ, Yacoob Y. Recognizing facial expressions in image sequences using local parameterized models of image motion. *Int J Computer Vis*. 1997;25(1):23–48.
4. Lucey S, Ashraf AB, Cohn JF. Investigating spontaneous facial action recognition through AAM representations of the face. In: Delac K, Grgic M, editors. *Face Recognition*. Vienna, Austria: I-Tech Education & Publishing; 2007. p. 275–86.
5. Horn BKP, Schunck BG. Determining optical flow. *Artif Intell*. 1981;17:185–203.
6. Tomasi C, Kanade T. Detection and Tracking of Point Features. Carnegie Mellon University; 1991.
7. Kanade T, Cohn J, , et al. Comprehensive database for facial expression analysis. In: *Proc IEEE Int Conf Automatic Face Gesture Recogn*; 2000. p. 46–53.

Kategorisierung der Beiträge

Modalität bzw. Datenmaterial

Röntgen, 66, 97, 147, 172, 201, 226, 241, 246, 251, 256, 271, 286, 296, 306, 315, 320, 325, 330, 335, 345, 370, 400, 405, 425, 440

- konventionell, 11, 16, 41, 51, 71, 76, 81, 86, 112, 117, 137, 142, 177, 196, 201, 206, 211, 221, 256, 261, 266, 276, 296, 301, 311, 325, 350, 385, 400, 410, 415, 435, 450
- digital, 71, 92, 117, 132, 276, 281, 335

Endoskopie, 16, 41, 51, 137, 206, 350, 410, 435

Optische Verfahren

- sonstige, 11, 16, 76, 81, 86, 112, 137, 177, 196, 211, 221, 261, 266, 301, 311, 385, 415, 450

Multimodale Daten, 16, 117, 142, 201, 256, 296, 325, 400

Angiographie, 107, 142, 162, 251

Computertomographie, 21, 31, 46, 107, 117, 122, 142, 157, 191, 216, 236, 256, 296, 325, 365, 380, 395, 400, 420, 430

- hochauflösend, 11, 61, 117, 375, 380, 400, 430
- spiral, 410, 430

Sonographie, 187, 355, 360, 390

Kernspintomographie, 97, 147, 172, 201, 226, 246, 251, 256, 271, 286, 325, 400, 405

- funktionell, 330
- hochauflösend, 66, 251, 306, 400, 425
- interventionell, 440

Positron-Emission-Tomographie, 201, 296, 325, 400

- hochauflösend, 201, 400

Single-Photon-Emission-Computertomographie, 241, 320, 345, 370

Szintigraphie, 335

Dimension der Daten

Signal (1D), 127

Bild (2D), 6, 16, 26, 51, 71, 92, 122, 137, 182, 196, 221, 241, 261, 266, 276, 281, 301, 315, 335, 345, 350, 365, 410, 415, 450

Bildsequenz (2D+t), 1, 41, 76, 112, 162, 206, 311, 385, 435, 450, 455

Volumen (3D), 11, 21, 36, 46, 56, 61, 66, 86, 97, 102, 107, 117, 132, 137, 142, 157, 162, 167, 177, 182, 187, 201, 211, 216, 226, 231, 236, 246, 251, 256, 271, 286, 291, 301, 306, 320, 325, 345, 355, 360, 370, 375, 380, 390, 395, 400, 405, 410, 420, 425, 430, 440, 445

Volumensequenz (3D+t), 16, 31, 81, 172, 291, 296, 330, 365, 385

Pixelwertigkeit

Einkanal, 1, 6, 11, 31, 66, 71, 81, 92, 107, 117, 122, 162, 172, 187, 191, 206, 211, 216, 221, 226, 236, 241, 256, 266, 271, 276, 281, 286, 291, 296, 301, 320, 325, 330, 335, 355, 360, 365, 375, 380, 390, 400, 405, 410, 425, 430, 455

Mehrkanal, 16, 26, 41, 51, 76, 86, 97, 102, 127, 132, 196, 221, 246, 261, 306, 311, 345, 350, 370, 385, 410, 415, 435, 450

Untersuchte Körperregionen

Ganzkörper, 137, 177, 226, 231, 425

Schädel, 46, 56, 102, 112, 117, 132, 142, 147, 162, 167, 182, 201, 221, 246, 256, 271, 286, 296, 301, 306, 330, 375, 395, 400, 425, 450, 455

Extremitäten

- obere, 61, 315, 335
- untere, 276, 335, 385

Thorax, 21, 31, 36, 102, 236, 281, 320, 365, 400, 410, 430, 440, 445

Mamma, 71, 92, 172, 390

Abdomen, 11, 16, 36, 46, 107, 137, 167, 206, 216, 350, 355, 405

Becken, 41, 51, 97, 117, 276, 335, 385, 435

Betrachtetes Organsystem

Systemübergreifend, 167, 196, 211, 226, 296, 311, 335, 375, 415, 445, 450

Immunzelluläres System, 196, 211, 311

Dermales System, 415

Zentrales Nervensystem, 147, 201, 246, 271, 286, 301, 306

Vegetatives Nervensystem, 261

- Kardiovaskuläres System, 36, 66, 107, 142, 162, 236, 320, 365, 425, 430, 440
- Respiratorisches System, 21, 31, 36, 152, 196, 410
- Gastrointestinales System, 16, 137, 206, 350, 355, 405
- Uropoetisches System, 41, 51
- Reproduktionssystem, 97
- Muskuloskeletales System, 61, 182, 276, 335, 385, 455
- Primärfunktion des Verfahrens**
- Bilderzeugung und -rekonstruktion, 1, 6, 16, 41, 51, 56, 102, 122, 127, 137, 182, 187, 201, 211, 221, 226, 241, 246, 256, 301, 320, 345, 370, 380, 440, 450
- Bildverbesserung und -darstellung, 11, 26, 41, 51, 102, 122, 157, 191, 206, 211, 226, 241, 246, 301, 335, 365, 390, 425, 435, 440, 445, 450
- Bildtransport und -speicherung, 132, 246, 251
- Merkmalsextraktion und Segmentierung, 66, 71, 76, 81, 86, 92, 97, 107, 132, 142, 172, 182, 187, 196, 221, 236, 246, 261, 266, 271, 276, 281, 286, 291, 296, 330, 350, 375, 385, 395, 405, 415, 420, 425, 430, 435, 455
- Objekterkennung und Szenenanalyse, 36, 81, 92, 117, 196, 231, 276, 281, 311, 315, 355, 375, 410, 455
- Quantifizierung von Bildinhalten, 61, 66, 76, 81, 86, 97, 112, 137, 191, 196, 211, 246, 276, 286, 296, 345, 360, 370, 375, 395, 415, 430, 435, 440
- Multimodale Aufbereitung, 16, 56, 127, 142, 162, 201, 325, 335, 400, 445
- Art des Projektes**
- Grundlagenforschung, 1, 6, 16, 61, 66, 76, 81, 137, 177, 196, 201, 241, 380, 415, 445
- Methodenentwicklung, 1, 6, 11, 26, 31, 36, 46, 56, 61, 71, 76, 81, 86, 92, 117, 122, 127, 142, 147, 157, 162, 167, 172, 182, 187, 191, 196, 201, 211, 221, 226, 231, 236, 246, 251, 256, 261, 266, 271, 276, 281, 291, 301, 306, 311, 315, 320, 335, 345, 350, 355, 365, 370, 375, 380, 385, 400, 405, 410, 415, 420, 425, 430, 435, 445, 455
- Anwendungsentwicklung, 36, 71, 97, 122, 152, 201, 206, 241, 271, 276, 286, 291, 296, 320, 325, 345, 350, 355, 360, 365, 390, 395, 430, 435
- Klinische Diagnostik, 21, 41, 51, 61, 66, 97, 102, 112, 226, 306, 320, 330, 335, 340, 415, 430, 440, 455

Autorenverzeichnis

- Aach T, 41, 51, 206, 320, 435
Adelt A, 177
Aichert A, 102
Althuizes M, 271
Ameling S, 350
Amunts CDK, 301
Athelougou M, 92
- Böhler T, 335
Backus B, 320
Baegert C, 216
Baer A, 167
Balda M, 177, **191**
Barthold C, **455**
Bartz D, 127, 142, 271
Becker S, **306**
Behrens A, **41**, **51**, 206, 435
Bendl R, 375
Beyerlein P, 281
Bichlmeier C, 102, 445
Biederer S, 1, **6**
Binnig G, 92
Bippus R, 296, 320, **370**
Bischoff B, 430
Bock R, 221
Boese J, 157
Bommers M, 41
Borgert J, 1
Born S, **142**
Botterweck H, 370
Boukamp P, 86
Braun A, 196
Broll M, 187
Bruijns J, 132
Buchfelder M, 147
Budai A, **261**
Buzug TM, 1, 6, 122, 286, 306
Bußmann S, 360
- Chen L, **132**
- Daum V, 221, 365
Dekomien C, 187
Delorme S, 226
Demirci S, 107
- Denecke T, 405
Dennerlein F, 157
Deserno TM, 251, 315, 340
Dey T, **320**
Dickhaus H, 415
Dinkel J, 226
Dornheim J, 420
Dornheim L, 395
dos Santos TR, 11, 21, **36**, 231
Drexl J, 390
- Eder M, 445
Ehrhardt J, 31
Eiben B, **301**
Eichelberg JRM, 340
Eils R, 81
Eisenbeis A, 330
Ellenberg J, 81
Ellenberg P, 92
Elter M, 71
Emmersberger M, **107**
Engel M, **216**
Engelbrecht R, 16
Engelhardt S, **350**
Englmeier KH, 196
Erfurt P, 56
Ernst C, 415
Essert-Villard C, 216
Euler E, 102
- Fangerau M, 216, 226
Feehan O, 92
Fiehler J, 330, 425
Fischer B, 315
Forkert ND, **330**, **425**
Fränzle A, 375
Franz AM, 231
Fried E, **251**
Friedl S, 455
Fritsche A, **315**
Fritzsche K, **246**
- Görtz A, 76
Gülpers R, 340
Günther RW, 340

- Gaens M, 201
 Gaspar M, 76
 Geng Y, 251
 Georgi J-C, **296**
 Gergel I, **21**, 36
 Ghotbi R, 107
 Gleich B, 1
 Godinez WJ, 81
 Goedicke A, 345, 370
 Gollmer ST, **286**
 Goößen A, **276**
 Grabe N, 415
 Greß O, **266**
 Greiner G, 147, 430
 Grigat R-R, 276
 Groch W-D, 117
 Gross S, 41, 51, **206**, 435
 Groth M, 66
 Grottke O, 251
 Grouls C, 340
 Guillem J, 296
 Gulbins (Grassmé) H, 211
 Gulbins E, 211
 Gurevich EL, 76
 Guski M, 51

 Hüttelmaier S, 266
 Hahn H, 390
 Hahn P, 395
 Hamm P, 196
 Hamo O, **256**
 Hamprecht FA, 97
 Handels H, 31, 66, 330, 425
 Harder N, **81**
 Hartmann P, 236
 Hausleiter J, 430
 Hautmann H, 410
 Heimann T, 236
 Heining SM, 26, 102
 Heismann BJ, 191
 Heldmann S, 335
 Hentschke CM, **162**
 Hermann E, 276
 Herzog A, 56
 Herzog H, 201
 Hofmann HG, **380**
 Hornegger J, 16, 152, 177, 191, 221, 261, 365, 380
 Humm JL, 296
 Hussong A, 56

 Ilgner J, **450**
 Illies T, 425
 Imhäuser C, **211**

 Jacob S, 137
 John C, **112**
 Joos S, 86
 Jung F, **325**

 Küblbeck TWC, 455
 Kaaks R, 226
 Kaffanke J, 201
 Kainberger F, 61
 Kassemeyer S, 97
 Kaster FO, **97**
 Keck B, 380
 Kellermann K, **167**
 Kirschner M, **291**
 Kneer D, 251
 Knopp T, **1**, 6
 Kops ER, 201
 Kovacs L, 445
 Krüger J, 385
 Kratz B, **122**
 Krefting D, 405
 Kuehne T, 440
 Kuhlen T, 251
 Kutter O, 102, 410

 Landes C, 182
 Landes J, 102
 Langs G, 61
 Lautenbacher S, 455
 Lee NY, 296
 Leonhardt S, 41
 Leypoldt F, 330
 Lichter P, 86
 Lipinski H-G, 211, 311
 Lippert H, 137
 Lorenz C, 281
 Lund G, 66

 Möller B, 266
 Möller D, 425
 Mühler K, **46**
 Müller H, 76, 182
 Müllerleile K, 66
 Maier-Hein L, 11, 16, 21, 36, 216, **231**
 Majdani O, 56
 Mang A, 306
 Martin-Gonzalez A, **26**

- Marwedel P, 76
 Mechtersheimer G, 86
 Meinzer H-P, 11, 16, 21, 36, 216, 226, 231, 236, 246, 440
 Mempel M, 196
 Menzel M, 410
 Merhof D, **147**
 Merkel B, 97
 Mersmann S, 11, 36
 Michaelis B, 137
 Michelson G, 261
 Mora-Bermúdez F, 81
 Mueller D, 61

 Narayanan M, 296
 Navab N, 26, 102, 107, 410, 445
 Nelles G, 256
 Nimsky C, 147
 Nix O, 97

 Oehler M, 122
 Onken M, 340
 Overhoff HM, 355, **360**

 Palm C, 301
 Papenberg N, **335**
 Papst A, 455
 Park J-HJ, 450
 Patsch J, 61
 Paulus D, 350
 Paulus J, **221**
 Penne J, 11, **16**, 152
 Pfannmöller M, 86
 Pfeifle M, 271
 Pietrzyk U, 301, 400
 Placht S, **177**
 Pommerencke T, **415**
 Posch S, 266
 Prümmer M, 365
 Pralow TGT, 276
 Preim B, 46, 167, 420
 Pritzkau A, **127**
 Prochiner AS, **355**
 Proksch D, **420**

 Rössling I, **395**
 Radmer J, **385**
 Rau TS, 56
 Redeleff BA, 216
 Reichl T, **410**

 Riefenstahl N, 137
 Rieker RJ, 86
 Riesenkampff E, 440
 Rietdorf U, **440**
 Ritschel K, 187
 Rohr K, 81, 86
 Romijn L, 320
 Rose G, 281
 Rossaint R, 251
 Ruppertshofen H, **281**

 Säring D, **66**, 330, 425
 Sakas G, 117
 Salah Z, 281
 Sander P, 86
 Sattel TF, 1, 6
 Schäfer S, **172**
 Schöder H, 296
 Schönmeier R, **92**
 Schaller C, 16, 152, 177
 Scheins J, 201
 Schink M, 206
 Schippritt D, **311**
 Schmauss B, 16
 Schmid U, 455
 Schmidt-Richberg A, 31, 425
 Schmidt G, 92
 Schnapauff D, 405
 Scholl I, **400**
 Schröder A, 182
 Schramm H, 281
 Schubert N, 400
 Schultis B, 410
 Schumacher H, 335
 Schwanecke U, 112
 Schwarz T, 226, 440
 Schweizer B, **345**
 Schwemmer C, **365**
 Seitel A, **11**, 216
 Selby BP, **117**
 Shah J, 201
 Siewert R, **405**
 Simbt S, **157**
 Sittek H, 92
 Sommer CM, 216
 Stöhr N, 266
 Stehle T, 41, 51, 206, **435**
 Stilla U, 117
 Stoll A, **375**

- Tönnies KD, 162, 172
Teßmann M, **430**
Tellmann L, 201
ter Romeny BMH, 132
Tetzlaff R, 11, 21
Timm C, 76
Tischendorf J, 206
Tolxdorff T, 405
Trautwein C, 206
- Ullrich S, 251
Ulrich C, **152**, 177
- Valentinitsch A, **61**
van Straaten D, 335
Vega-Higuera F, 430
Verzijlbergen J, 320
Voitel L, 271
- Wörmann J, **196**
Wünsche A, 81
Würfel W, **56**
Wagenknecht G, 256
Wagner F, **71**
Wagner M, 182
Walczak L, **182**
Wald D, **226**
Walter S, 117
Wang WW, 296
- Wang X, **236**
Wegner I, 21, 236
Weichert F, **76**, 182
Weirich C, **201**
Weizenecker J, 1
Wellein DI, 142, **271**
Welter P, **340**
Werner R, **31**
Wesarg S, 291, 325
Westhofen M, 450
Westphal K, 415
Wex C, 137
Wieczorek H, **241**, 320
Wieczorek M, **102**
Wiemann M, 311
Winter S, **187**
Wirth S, 350
Wirtz S, 335
Witte M, **137**
Wolf -J-C, 31
Wolf I, 236, 440
Wucherer P, **445**
Wulff J, 435
Wörz S, **86**
- Zöllner A, 142
Ziener P, 400
Zoehrer F, **390**
Zybin A, 76

Stichwortverzeichnis

- Abbildung, 301
- Ähnlichkeit, 1, 36, 92, 221, 236, 315
- Aktive Kontur, 271, 276
- Anatomie, 107, 276, 445
- Artefakt, 122
- Auflösung, 6, 61, 301
- Augmented Reality, 16, 26, 102, 410, 445
- Ausbildung, 46, 450
- Automat, 81

- B-Spline, 107
- Benutzerschnittstelle, 46
- Bewegung, 435
- Bewegungsanalyse, 31, 81, 172, 385, 435, 455
- Bewegungsunterdrückung, 21
- Bilddatenbank, 92, 251, 256
- Bildfusion, 16, 41, 51, 102, 162
- Bildgenerierung, 6, 102, 221, 226, 256, 320, 345
- Bildmontage, 41, 51
- Bildqualität, 1, 122, 157, 206, 241, 365, 390, 435
- Bildverbesserung, 16
- Bioinformatik, 81, 86, 127, 211, 415
- Biomechanik, 182, 276, 445

- CCD-Sensor, 301
- Clusteranalyse, 61
- Computer Aided Diagnosis (CAD), 41, 51, 61, 66, 71, 76, 92, 97, 132, 206, 340, 350, 430
- Computer Assisted Radiology (CAR), 216, 340
- Computer Assisted Surgery (CAS), 46, 132, 167, 216, 410
- Computer Based Training (CBT), 46, 137
- Content Based Image Retrieval (CBIR), 340
- Coocurrence Matrix, 350

- Datenbank, 251
- Datenreduktion, 296, 375
- Deformierbares Modell, 21, 86, 107, 276, 306
- Detektion, 172, 311, 350, 430
- Diffusion, 147, 211, 246, 390
- Digital Imaging and Communication in Medicine (DICOM), 340

- Echtzeit, 1, 6, 102, 167, 206, 380, 445
- Elastische Registrierung, 221
- Entwicklungsumgebung, 21, 256
- Entzerrung, 301
- Ergonomie, 450
- Erweiterte Realität, 410, 445
- Evaluierung, 11, 31, 97, 102, 157, 216, 256, 286, 320, 345, 365, 390, 395, 430
 - klinisch, 276, 296, 330
- Evolution, 187

- Filterung, 21, 261, 311, 365, 390, 415
 - nichtlinear, 196, 435
- Finite Elemente Modell (FEM), 182
- Fourier-Transformation, 97, 122
- Frequenzanalyse, 1, 76
- Fusion, 16

- Geometrie, 132, 395
- Gesichtsbildanalyse, 112, 455
- Gradient, 221
- Graph Matching, 36, 315
- Graphical User Interface (GUI), 246, 256

- Hardware, 16, 380
- Hochgeschwindigkeitskamera, 76
- Hough-Transformation, 281, 301, 360

- Image Retrieval, 92, 315
- Interface, 46, 167
- Interpolation, 122, 365

- Kalibrierung, 1, 16, 137, 157, 187
- Kantendetektion, 311, 405
- Klassifikation
 - statistisch, 71, 81, 97, 330, 375, 425
 - syntaktisch, 167, 405
- Klinische Evaluierung, 296, 450
- Komponentensoftware, 97, 246
- Kontur, 66, 107, 311
- Korrespondenzbestimmung, 36, 137, 286

- Kovarianzmatrix, 276
- Labeling, 97, 132, 261, 375
- Landmarke, 31, 276, 286, 455
- Lokalisation, 425, 430
- Marching Cube, 211
- Matching, 31, 36, 117, 335
- Minimalinvasive Chirurgie, 102, 137, 450
- Modellierung, 1, 137, 182, 256, 276, 296, 306, 360
- Morphologie, 196, 405
- Morphometrie, 286
- Multimedia, 46, 450
- Multiskalen, 51, 266
- Mutual Information, 117, 142, 221, 325
- Navigation, 21, 360, 410
- Numerik, 117, 182
- Oberfläche, 11, 107, 112, 137, 216, 286, 395
- Objekterkennung, 76, 81, 92, 196, 266, 281, 311, 375, 415, 425
- Objektverfolgung, 311
- Operationsplanung, 41, 167, 182, 216, 271
- Optimierung, 117, 157, 187, 296, 301, 380, 390
- Parametrierung, 66, 97, 167, 296
- Perfusion, 172, 320
- Picture Archiving and Communication System (PACS), 340
- Plattform, 92, 246
- Point-Distribution-Modell, 345
- Prädiktion, 435
- Qualitätskontrolle, 97
- Quantisierung, 66, 430
- Radon-Transformation, 117, 191
- Rauschen, 11, 191, 231, 241, 311, 390
- Region of Interest (ROI), 221, 380
- Region-Growing, 112, 132, 430
- Registrierung, 16, 31, 36, 112, 117, 142, 162, 221, 231, 256, 301, 325, 335, 410
 - elastisch, 221
- Regression, 385
- Rehabilitation, 385
- Rekonstruktion, 1, 191, 241, 370
 - 3D, 16, 112, 137, 157, 182, 211, 301, 320, 345, 365, 380, 425
- Rendering, 365
- ROC-Kurve, 76, 97
- Schwellwertverfahren, 66, 81, 167, 196, 311
- Screening, 430
- Simulation, 1, 6, 21, 117, 182, 251
- Skalenanalyse, 266
- Skalenraum, 76
- Spektralbild, 97
- Sprache, 137
- Strukturanalyse, 395
- Telemedizin, 450
- Textur, 61, 81
- Therapie, 46
- Therapieverlaufskontrolle, 112
- Tissue Engineering, 415
- Tracking, 81, 102, 147, 385, 410, 455
- Translation, 221
- Tumorstaging, 395
- Unified Modeling Language (UML), 97
- Validierung, 97, 256, 271, 345, 395, 430
- Video, 46, 206, 410, 445, 455
- Videodokumentation, 450
- Virtuelle Realität, 137, 251
- Visible Human, 102
- Visualisierung, 26, 46
 - 2D, 41, 51
 - 2D+t, 450
 - 3D, 46, 102, 107, 127, 137, 167, 211, 216, 246, 445
- Voice Control, 445
- Volume Rendering, 102, 117, 211, 410
- Volumetrie, 66
- Vorverarbeitung, 11, 206, 211, 226
- Warping, 26, 41, 51
- Wasserscheiden-Transformation, 196, 271, 415
- Wavelet-Transformation, 76, 206, 266, 350
- Workflow, 97
- Zeitreihe, 296, 311