

Erweiterte Realität und 3-D Visualisierung für minimal-invasive Operationen durch Einsatz eines optischen Trackingsystems

F. Vogt¹, S. Krüger², M. Winter³, H. Niemann¹, W. Hohenberger²,
G. Greiner³ und C.H. Schick²

¹Lehrstuhl f. Mustererkennung, Universität Erlangen-Nürnberg, 91058 Erlangen

²Chirurgische Universitätsklinik, 91054 Erlangen

³Lehrstuhl f. Graph. Datenverarbeitung, Universität Erlangen-Nürnberg, 91058 Erlangen

Email: vogt@informatik.uni-erlangen.de

Zusammenfassung. Optische Trackingsysteme bieten derzeit die größtmögliche Genauigkeit zur Positionsbestimmung eines bewegten Objektes. Es ist daher nahe liegend ein solches System zur Bestimmung der Endoskopposition in der minimal-invasiven Chirurgie einzusetzen, um den Operateur durch ein dreidimensionales Modell des Operationsgebietes (und damit die Möglichkeit der dreidimensionalen Betrachtung) sowie erweiterte Realität zu unterstützen. Der Einsatz von optischen Trackingsystemen für diese Zwecke wurde bisher nicht untersucht. Erstmals wird in diesem Beitrag eine Methode zur markerlosen Registrierung eines Endoskops mit CT/MR Daten vorgestellt. Die entwickelten Verfahren benötigen nur wenige Minuten Rechenzeit und sind daher für den Einsatz im Operationssaal geeignet.

1 Einleitung

Minimal-invasive Eingriffe werden in Körperhöhlen mit Hilfe spezieller chirurgischer Instrumente durchgeführt, wobei die Instrumente durch sehr kleine Zugänge (Trokare) eingeführt werden. Dadurch reduziert sich die Belastung für den Patienten erheblich, weshalb die Anzahl minimal-invasiver Eingriffe weiterhin zunimmt.

Da das Operationsgebiet nicht mehr direkt betrachtet werden kann, verwendet man Endoskope, die es ermöglichen das Operationsgebiet auf einem Videomonitor darzustellen. Die Belastung des Chirurgen ist allerdings im Vergleich zur herkömmlichen Operation deutlich erhöht. Gründe hierfür sind unter anderem die ungewohnte Betrachtung des Operationsgebietes auf einem Monitor, die dadurch reduzierte Tiefenwahrnehmung, die fehlende Haptik, die teilweise schlechte Bildqualität beispielsweise durch Rauchentwicklung beim Schneiden von Gewebe und das eingeschränkte Sichtfeld. Die Forschungen in diesem Bereich haben daher zum Ziel, die Belastung für den Chirurgen im Sinne einer computer-unterstützten Operation zu reduzieren.

In diesem Beitrag werden zwei Methoden zur Unterstützung bei minimal-invasiven Operationen vorgestellt. Mit Hilfe eines optischen Trackingsystems

wird aus den 2-D Bildern des Endoskops ein 3-D Modell des Operationsgebietes erzeugt. Als Modell wird das Lichtfeld [1] verwendet. Die beim offenen Eingriff vorhandene Tiefenwahrnehmung wird durch die Möglichkeit der dreidimensionalen Betrachtung des Lichtfeldes, beispielsweise auf einem 3-D Monitor, wieder hergestellt. Außerdem bietet das Lichtfeld die Möglichkeit das gesamte Operationsgebiet zu überblicken, was mit einem Endoskop nicht möglich ist. Desweiteren erlaubt das Lichtfeld die Bereitstellung von erweiterter Realität (Augmented Reality), indem die 3-D Information zur Registrierung mit CT/MR Daten verwendet wird. Die registrierten Daten können dann sowohl in das 2-D Live-Bild als auch in das Lichtfeld eingeblendet werden. Der Nutzen für den Operateur ergibt sich durch eine bessere Übersicht und die Vermeidung von Komplikationen, da der Verlauf von wichtigen aber nicht im Bild sichtbaren anatomischen Strukturen erkennbar wird. Beispielsweise sind dies Gefäße, die von anderen anatomischen Strukturen verdeckt werden.

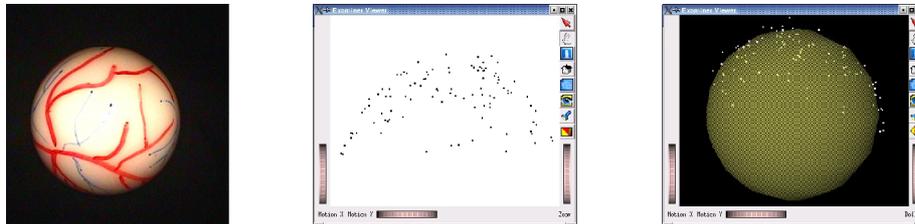
Die Erstellung dreidimensionaler Modelle für minimal-invasive Operationen ist Gegenstand aktueller Forschung [2, 3, 4]. Lediglich in [4] wird ein optisches Trackingsystem verwendet: zur Erstellung eines 3-D Panorambildes aus Endoskopbildern. Forschungsarbeiten zur Bereitstellung von erweiterter Realität existieren sowohl für herkömmliche als auch für minimal-invasive Operationen [5, 6]. Für erweiterte Realität bei minimal-invasiven Operationen wurden bisher magnetische Trackingsysteme und markerbasierte Registrierung eingesetzt [6]. Optische Trackingsysteme wurden in der Vergangenheit vor allem bei (offenen) neurochirurgischen Operationen eingesetzt.

2 Methoden

Für die Erstellung eines Lichtfeldes wird das Endoskop zunächst von Hand in beliebiger Abfolge über die Szene bewegt. Folgende Daten werden für qualitativ hochwertige Lichtfelder benötigt: intrinsische Kameraparameter (Brennweite, Hauptpunkt, Verzerrungsparameter), extrinsische Kameraparameter (Position und Orientierung) und Szenengeometrie für jedes aufgenommene Bild. Unter Kamera wird hierbei das gesamte optische System verstanden, bestehend aus Endoskop, Kamerakopf und Kamera. Die Position der gedachten Kamera befindet sich dabei im Projektionszentrum der Endoskopoptik (an der Spitze des Endoskops).

Die intrinsischen Kameraparameter werden durch Kamerakalibrierung bestimmt. Das optische Trackingsystem smARTtrack1 der Firma Advanced Realtime Tracking GmbH wird zur Bestimmung der extrinsischen Parameter verwendet. Hierzu wird ein Target benötigt, das am Endoskop befestigt wird. Da die Kameraposition und -orientierung nicht mit der Position und Orientierung des Targets übereinstimmen, muss zuvor die starre Transformation vom Target zur Kamera bestimmt werden. Dieses Problem ist als Hand-Auge-Kalibrierung bekannt und wird mit dem Algorithmus von Daniilidis [7] unter Zuhilfenahme eines Kalibrieremusters gelöst.

Abb. 1. Oberflächenrekonstruktion einer texturierten Kugel mit smARTtrack1: Originalbild (links), rekonstruierte 3-D Punkte der Kugeloberfläche (mitte) und geschätzte Kugel (rechts).



Szenengeometrie in Form der Szenenoberfläche wird rekonstruiert, indem markante Punkte von Bild zu Bild verfolgt und anschließend 3-D Punkte trianguliert werden. Resultat der Rekonstruktion ist eine Menge von 3-D Punkten, die sich auf der Szenenoberfläche befinden. Basierend auf der 2-D Nachbarschaft der projizierten 3-D Punkte wird für jedes Bild ein Dreiecksnetz der Oberfläche erstellt, das beim Rendern des Lichtfeldes verwendet wird.

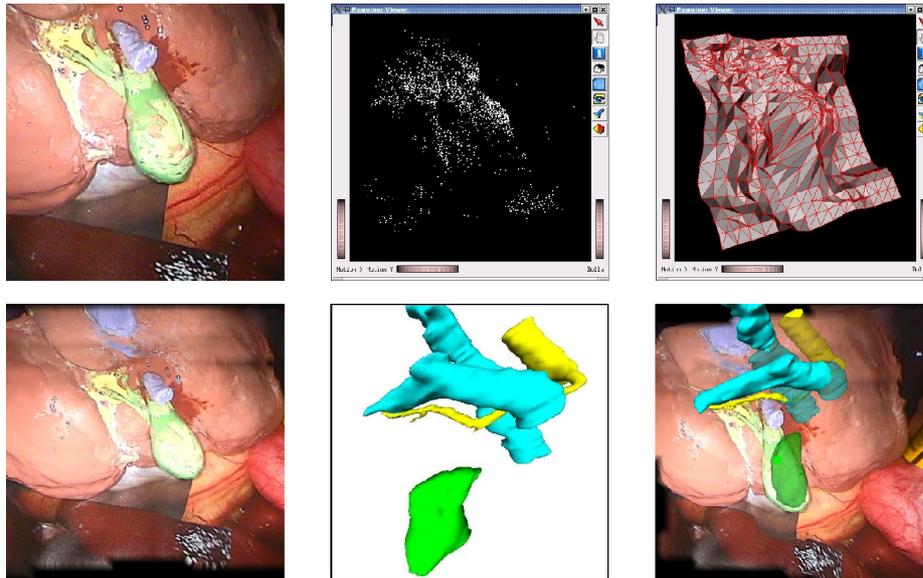
Die CT-Daten zur Registrierung werden von Hand segmentiert und mittels des „Marching Cube“-Isoflächen-Algorithmus [8] in Dreiecksnetze umgewandelt. Anschließend erfolgt die Registrierung halbautomatisch durch Auswahl von mindestens drei 3-D Punktkorrespondenzen im Lichtfeld und CT-Datensatz. Optional ist es möglich eine Feinregistrierung mit dem Iterative-Closest-Point (ICP) Algorithmus [9] durchzuführen. Abschließend können die segmentierten CT-Daten in korrekter Lage sowohl in das Live-Bild als auch in das Lichtfeld eingeblendet werden. Für MR-Daten ist der Ablauf analog.

3 Ergebnisse

Alle Versuche wurden mit einem OP-realistischen Versuchsaufbau durchgeführt, wobei eine laparoskopische OP-Situation (minimal-invasive Gallenblasenentfernung) simuliert wurde [2]. Die Hand-Auge Kalibrierung des Systems wurde mehrmals durchgeführt und deren Fehler bestimmt. Der mittlere Translationsfehler betrug 1,5 mm (4,3%), der mittlere Rotationsfehler betrug $0,2^\circ$ (1,2%).

Die Genauigkeit des Gesamtsystems wurde auf zweierlei Art bestimmt. Erstens wurde durch Aufnahme einer Sequenz eines Kalibrieremusters die Positions- und Orientierungsgenauigkeit bestimmt, indem die exakten Daten durch Kamerakalibrierung bestimmt und mit auf smARTtrack1-Daten basierenden Werten verglichen wurden. Der mittlere Translationsfehler des Gesamtsystems betrug 1,4 mm (4,5%), der mittlere Rotationsfehler betrug $0,3^\circ$ (4,8%). Zweitens wurde die Oberfläche einer Kugel mit einem Radius von 22,5 mm rekonstruiert (siehe Abb. 1). Die Abweichung betrug 1,4 mm (6%), wobei die Formabweichung der Kugel, d. h. die mittlere Abweichung der Punkte von der geschätzten Kugeloberfläche, 0,7 mm betrug.

Abb. 2. Erweiterte Realität: aus den Originalbildern (oben links) wird mit Hilfe von smARTtrack1 zunächst die Szenengeometrie in Form von 3-D Oberflächenpunkten (oben mitte) bzw. Dreiecksnetzen (oben rechts) rekonstruiert. Das Lichtfeld (unten links) wird mit den segmentierten CT-Daten (unten mitte) registriert und mit diesen überlagert (unten rechts). Es werden Gefäße im Bereich der Gallenblase überlagert. Die Originalsequenz wurde im Labor aufgenommen und zeigt ein Leber-Gallenblasen-Modell aus Silikon.



Mehrere Lichtfelder eines Leber-Gallenblasen-Modells aus Silikon wurden erstellt. Exemplarisch wurde eines dieser Lichtfelder anschließend mit segmentierten CT-Daten registriert. Dadurch konnte sowohl das Live-Bild wie auch das Lichtfeld mit CT-Daten überlagert werden (siehe Abb. 2). Die Erstellung eines Lichtfeldes aus 250 Bildern dauert dabei fünf Minuten (Punktverfolgung 3 min und Berechnung der Szenengeometrie 2 min). Die anschließende Registrierung erfordert eine weitere Minute.

4 Diskussion

Die vorgestellten Verfahren erlauben die schnelle Berechnung eines 3-D Modells des Operationsgebietes in Form eines Lichtfeldes bei minimal-invasiven Operationen. Darauf aufbauend kann durch Registrierung mit segmentierten CT-Daten erweiterte Realität zur Verfügung gestellt werden. Auf Grund der kurzen Rechenzeit ist der intraoperative Einsatz möglich und für Anfang des Jahres 2005 geplant.

Verglichen mit dem von unserer Arbeitsgruppe bisher eingesetzten Roboterarm [4] wurde insgesamt eine fünf Mal höhere Genauigkeit bei der Bestimmung

der extrinsischen Kameraparameter erreicht. Der im Vergleich zum Rotationsfehler größere Translationsfehler ergibt sich aus der Tatsache, dass es für die Rotation unerheblich ist, wo sich das Endoskop relativ zum Target befindet. Der Fehler liegt im vom Hersteller angegebenen Bereich. Bei der Translation verschlechtert die Entfernung des Targets vom Kamerazentrum (Endoskopspitze) die Genauigkeit. Da sich die Endoskopspitze im Patienten befindet, lässt sich die Entfernung nicht verringern. Während einer Operation beträgt der Abstand ca. 40 cm, somit ergibt sich bei einem Rotationsfehler von $0,2^\circ$ am Target bereits eine Abweichung von 1 mm an der Endoskopspitze.

Nachteile des optischen Trackingsystems resultieren lediglich aus der erforderlichen Sichtbarkeit des Targets und der Verdeckungen von Target-Markern, wobei letzteres durch ein geeignetes Target-Design weitgehend vermieden werden kann. Magnetische Trackingsysteme oder endoskopführende Roboterarme sind in dieser Hinsicht flexibler, dafür kann die Position und Orientierung aber nicht so genau bestimmt werden. Die Qualität der vorgestellten Verfahren hängt allerdings entscheidend von der erreichten Genauigkeit ab.

Diese Arbeit wurde durch die DFG im Rahmen des SFB 603 (TP B6) gefördert. Für den Inhalt sind ausschließlich die Autoren verantwortlich.

Literaturverzeichnis

1. Steven JG, Grzeszczuk R, Szeliski R, Cohen MF. The Lumigraph. In: Computer Graphics Procs, Annual Conference Series (Procs SIGGRAPH '96); 1996. p. 43–54.
2. Thormählen T, Broszio H, Meier PN. Automatische 3D-Rekonstruktion aus endoskopischen Bildfolgen. In: Procs BVM; 2002. p. 207–210.
3. Vogt F, Krüger S, Schmidt J, Paulus D, Niemann H, Hohenberger W, et al. Light Fields for Minimal Invasive Surgery Using an Endoscope Positioning Robot. *Methods of Information in Medicine* 2004;43(4):403–408.
4. Dey D, Gobbi DG, Slomka PJ, Surry KJM, Peters TM. Automatic Fusion of Free-hand Endoscopic Brain Images to Three-Dimensional Surfaces: Creating Stereoscopic Panoramas. *IEEE Transactions on Medical Imaging* 2002;21(1):23–30.
5. Vogt S, Wacker F, Khamene A, Elgort DR, Sielhorst T, Niemann H, et al. Augmented Reality System for MR-guided Interventions: Phantom Studies and First Animal Test. In: Procs of SPIE's Conference of Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display; 2004. p. 100–109.
6. Scheuring M. Fusion of Medical Video Images and Tomographic Volumes. Ph.D. thesis. Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany; 2003.
7. Daniilidis K. Hand-Eye Calibration Using Dual Quaternions. *International Journal of Robotics Research* 1999;18:286–298.
8. Lorensen WE, Cline HE. Marching cubes: A high resolution 3D surface construction algorithm. In: Procs of the 14th annual conference on Computer graphics and interactive techniques. ACM Press; 1987. p. 163–169.
9. Besl PJ, McKay ND. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1992;14(2):239–256.