

Strukturprototypen zur Modellierung medizinischer Bildinhalte

Benedikt Fischer, Benjamin Winkler, Christian Thies,
Mark O. Güld und Thomas M. Lehmann

Institut für Medizinische Informatik, RWTH Aachen, 52057 Aachen
Email: bfischer@mi.rwth-aachen.de

Zusammenfassung. Medizinische Bildinhalte sind häufig aus mehreren Bildregionen zusammengesetzt, die zueinander in bestimmten Verhältnissen stehen. Diese lassen sich über Wahrscheinlichkeitsverteilungen regionaler und relationaler Attribute zu einem Strukturprototypen verallgemeinern und mit Szenen unbekanntem Inhalts vergleichen. Die semi-automatische Synthese eines Strukturprototyps sowie seine Anwendung zur Szenenanalyse wird am Beispiel von Handradiographien vorgestellt. Durch Verwendung der Strukturinformation wird hier eine Steigerung der Erkennungsrate um 17% erzielt.

1 Einleitung

Aufgrund der bei medizinischem Bildmaterial auftretenden hohen Variabilitäten lassen sich nur schwer Normwerte für die Attribute zu identifizierender Objekte entwickeln. So sind beispielsweise die Regionenmerkmale von Radiographien der Metakarpalknochen nicht ausreichend trennscharf, um eine korrekte Zuordnung der Regionen zu den einzelnen Fingern zu erlauben. Auch für einen menschlichen Betrachter ist es nahezu unmöglich, einen einzelnen Knochen ohne Umgebung korrekt zu klassifizieren. Dies fällt jedoch leicht, wenn die Verhältnisse zu anderen Knochen (Position, Größe, etc.) bekannt sind.

Eine Modellierung einer solchen Szene kann über einen attributierten relationalen Graphen (ARG) erfolgen, der ausschließlich die relevanten Regionen als Knoten und ihre Verhältnisse zueinander als Relationskanten repräsentiert. Diese Szene-ARGs kommen daher mit nur wenigen Elementen aus.

Bei der Segmentierung von Bildern unbekanntem Inhalts müssen jedoch wegen der unbekanntem Auflösung je Pixel unterschiedliche Regionenzugehörigkeiten betrachtet werden. In [1] wird das Bild dazu in einem iterativen Regionenverschmelzungsprozess in einen hierarchischen attributierten Regionenadjazenzgraphen (HARAG) überführt. Die resultierenden Graphen sind dadurch jedoch deutlich größer als die Szene-ARGs.

Eine Szenenanalyse entspricht dem Vergleich des Szene-ARG mit dem HARAG des zu analysierenden Bildes. Zur allgemeinen Objektsuche anhand von ARGs existiert in der Fachliteratur eine Reihe von Lösungsansätzen [2, 3]. Oft basieren die Referenzen jedoch lediglich auf manuell erzeugten Segmentierungen einer Referenzaufnahme oder auf Kantenbildern, deren Anwendung wegen

der Delineationsproblematik ausscheidet. Häufig können mit den vorgestellten Lösungen auch lediglich kleine Graphen mit weniger als 100 Knoten effizient miteinander verglichen werden oder die hohen Variabilitäten in den relevanten Bildstrukturen werden nur unzureichend modelliert. Zudem können lediglich Graphen gleichen Typs miteinander verglichen werden.

2 Methoden

Die Strukturprototypgraphen verallgemeinern die relevanten Bildinhalte für einen Anwendungsfall, so dass Variabilitäten nicht nur in Einzelobjekten sondern auch im Zusammenhang berücksichtigt werden. Die Repräsentation, semiautomatische Synthese und Anwendung des Prototypgraphens wird im Folgenden vorgestellt. Als Anwendungsbeispiel werden Skelettradiographien der linken Hand herangezogen, wie sie bei Kindern zur Bestimmung des Knochenalters verwendet werden.

2.1 Prototyprepräsentation und -training

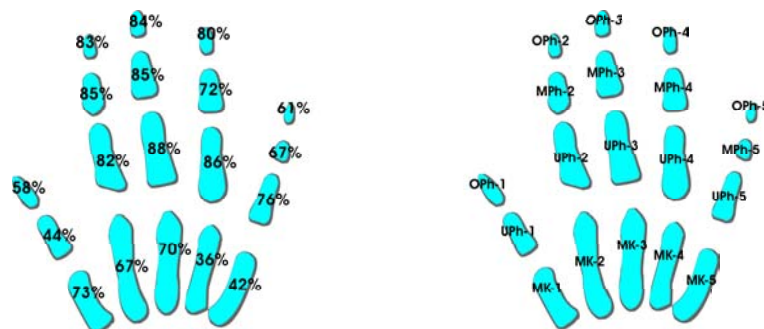
Der Prototypgraph ist ein vollständiger Graph, dessen Knoten Einzelobjekte und dessen Kanten Relationen zwischen den Einzelobjekten beschreiben. Im Anwendungsbeispiel entsprechen die Einzelobjekte den einzelnen Handknochen. Für jedes Einzelobjekt werden die statistischen Verteilungen der regionalen Merkmale als Knotenattribute, die der relationalen Merkmale zwischen den Einzelobjekten als Kantenattribute kodiert. Beispiele für ein regionales bzw. relationales Merkmal sind der mittlere Grauwert einer Region bzw. die relative Größe zweier Regionen. Die Merkmale des Prototypgraphens sind dabei unabhängig von den Merkmalen der zur Segmentierung verwendeten HARAGs. Insbesondere werden die relationalen Merkmale erweitert.

Die Merkmalsverteilungen werden in einer Trainingsphase unter Annahme von Normalverteilungen geschätzt: Als Referenz werden zunächst die relevanten Knoten für jedes Einzelobjekt in automatisch erzeugten HARAGs manuell mit einem entsprechenden Benutzerinterface identifiziert. Die bereits während der Segmentierung berechneten regionalen Merkmale werden als Knotenattribute weiterverwendet. Bei Bedarf können auch weitere regionale Merkmale berechnet werden. Danach werden je Datensatz Kanten zwischen den identifizierten Knoten eingefügt und mit relationalen Merkmalen attribuiert. Anschließend kann dann die Verteilung aller Attribute geschätzt werden. Insgesamt wurden 34 regionale Merkmale zu Form, Intensität und Textur sowie 12 relationale Merkmale zu Topologie-, Größen- und Intensitätsverhältnissen verwendet.

Mit Hilfe der ermittelten Verteilungen wird dann eine Ähnlichkeitsfunktion für jeden Prototypknoten bzw. jede Prototypkanten definiert, um die Ähnlichkeit unbekannter Regionen bzw. Regionenpaare zu den Einzelobjekten bzw. Einzelobjektpaaren zu bestimmen.

Im Anschluss an die Verteilungsberechnung findet eine Merkmalsselektion statt, um die besten Merkmalsmengen je Region und Relation zu bestimmen und

Abb. 1. Prozentuales Auftreten der einzelnen Handknochen als eigenständige Region (links) und Nummerierung der Prototypklassen (rechts). Die Nummerierung besteht aus der Knochenklasse (MK: Metakarpal, Uph: untere, MPh: mittlere, bzw. OPh: obere Phalanx) sowie der Finger Nummer (1:Daumen bis 5:kleiner Finger).



den Lösungsraum für das Graph-Matching zu verkleinern. Im Anwendungsbeispiel wurde mit *Sequential Forward Selection* ein ebenso einfaches wie schnelles Verfahren ausgewählt [4].

2.2 Vergleich Prototyp- und Bildgraph

Der Vergleich zwischen Prototyp- und Bildgraph erfolgt auf Basis eines Hopfield-Netzes [5]. In einem Assoziationsgraphen werden Entsprechungshypothesen zwischen Prototypknoten und HARAG-Knoten als eigener Graphknoten repräsentiert, sofern eine grundlegende Ähnlichkeit der repräsentierten Regionen über die regionalen Merkmale vorliegt. Diese dient dann auch gleichzeitig als initiales Potential der Knoten. Das Verfahren wurde modifiziert, um die relationalen Merkmale des Prototypen zur Stärkung bzw. Hemmung von Hypothesen ebenfalls berücksichtigen zu können. Für die Matching-Hypothesen werden zwischen betrachteten HARAG-Knoten im Bedarfsfall fehlende relationale Merkmale berechnet. Über ein Gradientenabstiegsverfahren werden die Potentiale der Matching-Hypothesen solange verändert, bis die gefundene Zuordnung stabil ist.

2.3 Experimente

Zur Evaluierung wurden 96 Handradiographien verwendet. Sofern in der Segmentierung vorhanden, wurden je Bild bzw. Graph 19 Handknochen (Metakarpalknochen bis obere Phalanx) mit einem Klassenlabel versehen (Abb. 1). Insgesamt wurden 221614 Regionen segmentiert, von denen 1290 den 19 Klassen entsprechen. Abb. 1 illustriert, wie oft die jeweiligen Knochen in der Segmentierung vorhanden waren. Durchschnittlich bestand ein HARAG aus 2300 Knoten.

Tabelle 1. Erkennungsrate der einzelnen Fingerknochen

Klasse	vorh.	lokal	strukturell	Klasse	vorh.	lokal	strukturell
MK-1	70	47	52	MPh-2	82	45	53
MK-2	64	45	50	MPh-3	82	59	68
MK-3	67	46	53	MPh-4	72	52	63
MK-4	35	24	26	MPh-5	64	38	43
MK-5	40	16	27	Oph-1	56	21	32
UPh-1	42	27	29	Oph-2	80	28	43
UPh-2	79	58	59	Oph-3	81	44	44
UPh-3	84	59	63	Oph-4	77	36	47
UPh-4	83	67	71	Oph-5	59	21	26
UPh-5	73	43	60				
Gesamt					1290	776	909
					100%	60.2%	70.5%

Zur Evaluierung wurde eine sechsfache Kreuzvalidierung gewählt, um jeweils einen Prototypen zu berechnen und die Erkennungsleistung zu bestimmen. Eine Regionenzuordnung wird als korrekt gezählt, wenn sich die vom Prototypen zugeordnete Region mit der Segmentierung der Ground Truth zu mindestens 85% überlappt. Eine hundertprozentige Überdeckung wurde vernachlässigt, da sich Regionen oft nur in wenigen kleinen Details unterscheiden und auch manuell kaum zu beurteilen ist, welche Region „die“ korrekte Lösung darstellt. So werden zum Beispiel bei manchen Regionen die Kapseln in die Segmentierung des Fingers einbezogen, in anderen Fällen dagegen nicht.

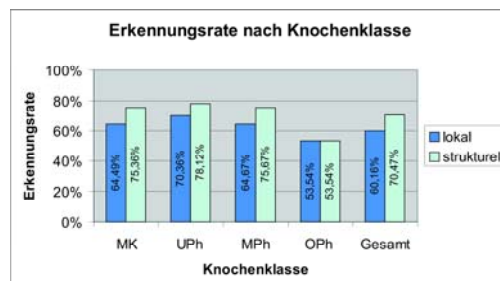
3 Ergebnisse

Über alle Klassen gemittelt wurden ohne relationale Merkmale lediglich 60.2% der Regionen korrekt klassifiziert, mit Hilfe des Strukturprototypen konnte dagegen eine Erkennungsrate von 70.5%, d.h. eine Steigerung um über 17%, erzielt werden (Tab. 1). Abbildung 2 gruppiert die Resultate über die jeweilige Knochenklasse. Bei allen Klassen wird eine Verbesserung durch Verwendung der Strukturinformationen erreicht. Das Trainieren der sechs Prototypen beanspruchte durchschnittlich 15 Minuten, das jeweilige Testen der 16 Datensätze dagegen lediglich 30 Sekunden.

4 Diskussion und Ausblick

Die durchgeführten Experimente illustrieren die generelle Anwendbarkeit von Strukturprototypen zur Objektsuche. Die Strukturprototypen führen in allen Belangen zu einer deutlichen Verbesserung der Erkennungsrate gegenüber Vergleichen ohne strukturelle Information. In Abbildung 2 fällt jedoch das schlechte Abschneiden der Klasse OPh auf. Dies ist dadurch zu erklären, dass die Regionen dieser Klasse sehr klein sind, die Regionengröße jedoch einen starken Einfluss auf

Abb. 2. Vergleich von lokaler Suche und Struktursuche mit einer geforderten Regionenüberdeckung von 85%.



viele der verwendeten Merkmale hat. So wirken sich z.B. kleine Änderungen der Regionen deutlich auf Formmerkmale aus. Zudem existieren in den HARAGs bedingt durch den Verschmelzungsprozess viele kleine Regionen, so dass hier eine Abgrenzung der richtigen Regionen schwieriger ist. Neben den hierdurch zu erklärenden schlechten Werten für OPh-1 und OPh-5 stehen in Tabelle 1 auch die Klassen MK-4, MK-5 sowie UPH-1 negativ hervor. Wie aus Abb. 1 hervorgeht, standen genau für diese Klassen deutlich weniger Trainingsdaten zur Approximation der Merkmalsverteilungen zur Verfügung als für die restlichen Klassen, so dass der Grund für die geringe Erkennungsleistung hierin zu sehen ist.

Abhilfe soll deswegen in Zukunft durch eine robustere Segmentierung geschaffen werden. Bislang wurde außerdem lediglich eine Normalverteilung der verwendeten Merkmale angenommen, statt die Verteilungen genauer zu schätzen. Zusammen mit optimierten Trainingsmethoden für die Merkmalsgewichtung, z.B. durch eine Supportvektormaschine, bieten sich hierdurch weitere Verbesserungsmöglichkeiten für künftige Forschungsarbeiten.

Literaturverzeichnis

1. Lehmann TM, Beier D, Thies C, Seidl T. Segmentation of medical images combining local, regional, global, and hierarchical distances into a bottom-up region merging scheme. In: Procs SPIE. vol. 5747; 2005. p. 545–555.
2. Petrakis GM. Design and evaluation of spatial similarity approaches for image retrieval. Image and Vision Computing 2002;20(1):59–76.
3. Park BG, Lee KM, Lee SU, Lee JH. Recognition of partially occluded objects using probabilistic ARG (attributed relational graph)-based matching. Computer Vision and Image Understanding 2003;90:217–241.
4. Cantu-Paz E, Newsam S, Kamath C. Feature selection in scientific applications. In: Procs ACM International Conference on Knowledge Discovery and Data Mining; 2004. p. 788–793.
5. Lappe C, Fischer B, Thies C, Güld MO, Kohnen M, Lehmann TM. Optimierung eines konnektionistischen Graphmatchers zum inhaltsbasierten Retrieval medizinischer Bilder. In: Procs BVM; 2004. p. 338–342.