Evolutionäres Graphmatching zur Handknochen-Identifikation

Benedikt Fischer, Armin Fritsche, Christian Thies, Thomas M. Deserno

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

Kurzfassung. Im medizinischen Umfeld ist häufig eine Identifikation lokaler Bildinhalte (z.B. Organe) erforderlich. Dazu lassen sich Bildinhalte als Graphstruktur abstrahieren und mit für die jeweilige Anwendung statistisch ermittelten Strukturprototypen vergleichen. Die Identifikation unbekannter Inhalte zu den Prototypen entspricht dann einem Graphmatching. In dieser Arbeit wird dazu ein auf evolutionärer Spieletheorie basierender Ansatz verwendet. Die Praxistauglichkeit wird am Beispiel der Identifikation von 19 Handknochen in 96 Handradiographien evaluiert und mit weiteren Matchingverfahren verglichen. Bei durchschnittlich 2.300 hierarchisch segmentierten Regionen je Radiographie werden mit 34 regionalen und 12 relationalen Merkmalen Werte von bis zu 71,63%, 65,76% bzw. 64,08% für Recall, Precision bzw. F-Measure für 27 unterschiedliche Parametrierungen erzielt.

1 Einleitung

Zur Identifikation medizinischer Bildinhalte werden häufig Graphen verwendet, da diese eine abstrakte relationale Strukturbeschreibung erlauben [1, 2, 3]. Dies ist insbesondere hilfreich, wenn eine punktweise Registrierung wenig erfolgversprechend ist, z.B. bei multimodalen Bilddaten mit unterschiedlicher Semantik der Bildkanäle, verschiedenen Skalierungen, oder auch Verformung und Verschiebungen beispielsweise durch Atmung oder Bewegung. Der Vergleich der Graphen wird über Graphmatching-Verfahren durchgeführt, wie sie auch in vielen Anwendungsdisziplinen außerhalb der Medizin eingesetzt werden [4].

Um die für eine Anwendung typischen Strukturen zu generalisieren, lassen sich Strukturprototypen definieren [3]. Diese repräsentieren Objekte im Bild als Graphknoten und modellieren die anwendungstypischen Eigenschaften über die statistische Verteilung regionaler und relationaler Merkmale. Zu registrierende Bilder werden zunächst in einen hierarchischen attributierten Regionenadjazenzgraph (HARAG) überführt, der dann über Graphmatchingverfahren mit Prototypen verglichen wird (Abb. 1). Die in [3] verwendeten Graphmatching-Verfahren sind jedoch mit teils umfangreichen Parametersätzen verbunden. Mit [5] ist dagegen ein evolutionärer Graphmatcher für das Matching von Gefäßbäumen bekannt, der lediglich wenige Parameter erfordert [1]. Allerdings werden in [5] zwei identische Graphtypen verglichen, welche die Prototypabstraktion und die generalisierende Struktur des HARAG unberücksichtigt lassen. Im Rahmen dieser

Arbeit wird deshalb der evolutionäre Graphmatcher für diese Zwecke angepasst und exemplarisch evaluiert.

2 Material und Methoden

Zum Vergleich der Graphen wird oft auf Assoziationsgraphen zurückgegriffen, bei denen jeder Graphknoten einer Zuordnungshypothese zwischen zwei Knoten aus den zu vergleichenden Graphen entspricht [3, 5]. Einzelne Hypothesen können durch Kanten bestärkt oder gehemmt werden. Die Suche nach dem optimalen Matching entspricht der nach einer maximalen Clique, wofür die Methodik aus [6] zum Einsatz kommt. Diese kann jedoch nicht direkt auf den Assoziationsgraphen für Prototypen und HARAGs angewendet werden. Zum einen operiert das Verfahren aus [6] nur auf kantengewichteten Graphen ohne Knotengewichtung, wodurch sämtliche regionalen Merkmale entfallen würden. In dieser Arbeit werden die regionalen Merkmale daher mit in die Kantengewichtung integriert. Zum anderen werden in [6] für die Verwendung eines Strafterms bei Mehrfachzuordnungen positive Kantengewichte voraussetzt. In dieser Arbeit wird statt dessen ein Benefit eingesetzt, der eine Unterscheidung zwischen unähnlichen Relationen (d.h. Kantenmerkmalen) und sich widersprechenden Hypothesen gewährleistet. Der entsprechend angepasste Assoziationsgraph ist wie folgt definiert:

Definition 1. (*angepasster Assoziationsgraph)*
Für zwei Graphen $G = (V_G, E_G)$ *und* $H = (V_H, E_H)$ *mit*

- *Knotenmarkierung* $\ell_G(x_i)$ *für* $x_i \in V_G$ *bzw.* $\ell_H(y_k)$ *für* $y_k \in V_H$,
- *Kantenmarkierung* $\ell_G(x_i, x_j)$ *für* $x_i, x_j \in V_G$ *bzw.* $\ell_H(y_k, y_l)$, *für* $y_k, y_l \in V_H$,
- *Ähnlichkeitsfunktionen* σ_V *bzw.* σ_E *für die Knoten- bzw. Kantenähnlichkeit,*
- *Schwellwerten* θ_V *bzw.* θ_E *für* σ_V *bzw.* σ_E

wird der angepasste Assoziationsgraph $A = (V_A, E_A)$ *wie folgt konstruiert:*

- $V_A = \{(x_i, y_k) : x_i \in V_G \text{ und } y_k \in V_H \text{ und } \sigma_V(\ell_G(x_i), \ell_H(y_k)) \geq \theta_V\}$,
  d.h. für jedes Knotenpaar der Ausgangsgraphen mit hinreichender regionaler Ähnlichkeit wird ein Hypothesenknoten erzeugt.
- *Die Knoten werden nicht markiert.*
- $((x_i, y_k), (x_j, y_l)) \in E_A$ *gdw.* $i \neq j$ *und* $k \neq l$
  d.h. zwischen allen nicht widersprüchlichen Matchinghypothesen wird eine Kante erzeugt.
- $\ell_A((x_i, y_k), (x_j, y_l)) = \begin{cases} ben + \frac{s_{i,k}+s_{j,l}}{2} + \frac{s_{ij,kl}}{2} & falls \frac{s_{ij,kl}}{2} > \theta_E \\ ben & falls \frac{s_{ij,kl}}{2} \leq \theta_E \end{cases}$
  d.h. alle hinreichend ähnlichen Kanten werden mit der angepassten Kantengewichtung markiert. Unähnliche, aber nicht in Widerspruch zu anderen Hypothesen stehende Kanten erhalten eine Markierung mit der Benefitgewichtung.

Zusammen mit der in [3] über eine gewichtete Mahalanobis-Distanz definierten Ähnlichkeit wird die angepasste Kantengewichtung für nichtnegative Kanten

aus der lokalen Ähnlichkeit $s_{i,k}$, der relationalen Ähnlichkeit $s_{ij,kl}$ der beiden Hypothesen sowie einem Benefitwert *ben* gebildet:

$$s_{i,k} = \sigma_V(\ell_G(x_i), \ell_H(y_k)) \tag{1}$$

$$s_{ij,kl} = \sigma_E(\ell_G(x_i, x_j), \ell_H(y_k, y_l)) + \sigma_E(\ell_G(x_j, x_i), \ell_H(y_l, y_k)) \tag{2}$$

Zum einen lassen sich auf diese Weise die Knotengewichtungen in Kantengewichtungen überführen. Zum anderen können über *ben* die sich nicht widersprechenden Zuordnungshypothesen belohnt werden. Die Adjazenzmatrix des wie obig konstruierten Assoziationsgraphen bietet nun eine korrekte Eingabe für das Verfahren aus [6]. Neben dem obig eingeführten Parameter *ben* müssen außerdem zwei für den Evolutionsabbruch erforderliche Parameter gesetzt werden. Zunächst wird über den so genannten ϵ-Support [1] geprüft, ob die Bewertungen aller nicht-ausgestorbenen Hypothesen größer als ein Parameter ϵ sind. Weiter wird über eine Variable Δ_ϵ beobachtet, ob sich der ϵ-Support über eine bestimmte Anzahl von Generationen nicht mehr ändert. Aus [7] ist bekannt, dass die Größe des ϵ-Supports monoton fallend ist, wenn ϵ klein genug gewählt wurde.

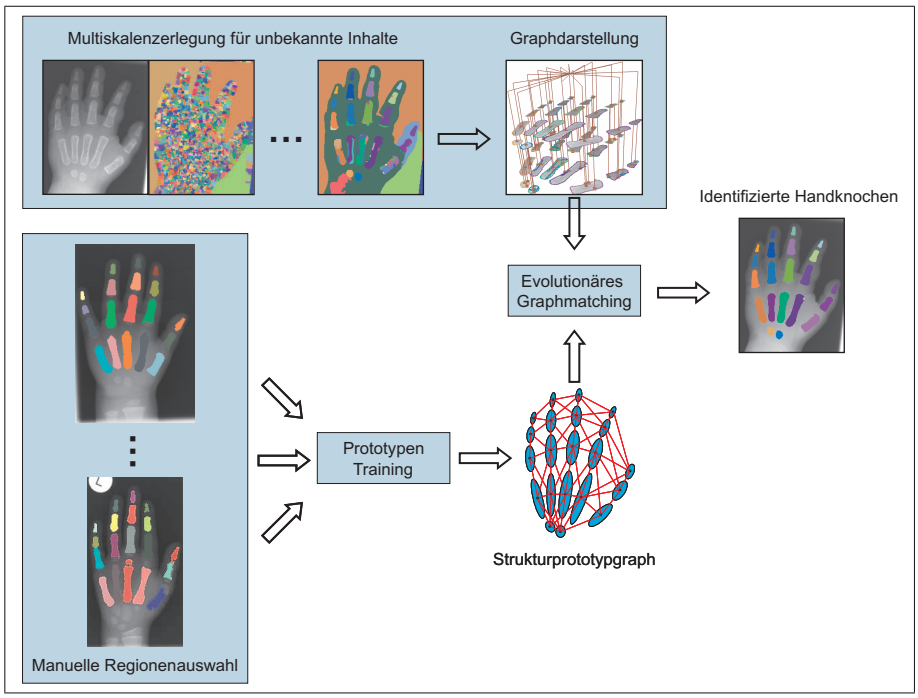

Abb. 1. Prinzip der Handknochenidentifkation: Bilder unbekannten Inhalts werden durch eine hierarchische Segmentierung in Graphen überführt. Diese werden mit Strukturprototypgraphen, die aus manueller Regionenauswahl synthetisiert werden, über den evolutionären Graphmatcher verglichen und die relevanten Regionen als Handknochen identifiziert.

Ist ϵ zu groß gewählt, könnte eine Hypothese die nur kurzzeitig geschwächt wird (etwa bis eine andere Hypothese ausstirbt) zu früh als ausgestorben definiert werden. Die Wahl Δ_ϵ hat ebenfalls Einfluss auf das Verfahren. Ist der Wert zu klein gewählt, ist es möglich, dass später noch weitere Hypothesen aussterben. Ist der Wert zu hoch gewählt, so kann es „nur“ zu längeren Laufzeit führen.

Die Praxistauglichkeit des evolutionären Matchings zur Objektidentifikation wird anhand von Experimenten zur Identifizierung von Handknochen experimentell ermittelt. Dazu wurden 96 Radiographien von linken Händen in sechsfacher Kreuzvalidierung mit Strukturprototypen verglichen. Zur Validierung liegt zu jedem Matching eine von Hand erstellte Groundtruth vor. Ein Matching für eine Region wird als korrekt gewertet, wenn die zugeordnete Region mit der Region aus der Groundtruth zu mindestens 85% überlappt. Diese Einschränkung ist erforderlich, da durch die hierarchische Segmentierung teilweise mehrere übereinander liegende Regionen als korrekt angesehen werden können, allerdings nur eine davon in der Groundtruth markiert ist. Zur Evaluation werden die Standardwerte *Recall*, *Precision* und *F-Measure* sowie die Laufzeiten ausgewertet. Die Laufzeiten wurden auf einem Linux Rechner mit 2.8 GHz Intel Xeon Prozessor und 2 GB RAM ermittelt. Die Berechnung erfolgte mit jeweils drei Parameterwerten ($\{1, 2, 3\}$, $\{0, 10^{-5}, 10^{-2}\}$, bzw. $\{25, 50, 100\}$) für die drei Parameter *ben*, ϵ, und Δ_ϵ, so dass insgesamt 27 Parameterkombinationen getestet wurden. Um die Ergebnisse beurteilen zu können, werden sie mit den jeweiligen Werten der beiden besten Verfahren aus [3] verglichen.

3 Ergebnisse

Die verwendeten Bildgraphen enthalten jeweils zwischen 1.000 und 6.000 Knoten, durchschnittlich 2.300 und insgesamt 221.615 Knoten. Die jeweils aus den Trainingsdaten der Kreuzvalidierung erzeugten Handstrukturprototypen besitzen 19 Knoten zur Repräsentation der Phalanx- und Metakarpalknochen. Über alle getesteten Parametereinstellungen gemittelt ergeben sich für den evolutionären Matcher 60,37%, 60,63% und 60,49%, bzw. 1,17s für Recall, Precision und F-Measure bzw. für die mittlere Laufzeit. Die besten erzielten Werte mit jeweils eigenen Parametersätzen für Recall, Precision, F-Measure bzw. Laufzeit sind 71,63%, 65,76%, 64,08% bzw. 0,46 Sekunden. Die besten Ergebnisse hinsichtlich Recall im Vergleich zu den beiden besten Verfahren aus [3] sind in Tabelle 1 zusammengefasst.

4 Diskussion

Das hier untersuchte Verfahren stellt in allen Belangen eine Verbesserung zu den bisher eingesetzten Verfahren aus [3] dar. Wie Tabelle 1 belegt, konnte der Recall um 1,6% bzw. 2,6% in Bezug auf den Hopfield-Matcher bzw. Similarity Flooding gesteigert werden. Noch deutlicher ist der gleichzeitige Zugewinn der Precision (18,9% bzw. 10,4%), F-Measure (11,9% bzw. 7%) und Laufzeit (Reduktion um mindestens 78,9% bzw. 12%). Sofern die Parameter nicht wie in Tabelle 1 auf

Tabelle 1. Ergebnisse des Matchingverfahrens im Vergleich mit den beiden in Bezug auf Recall besten Verfahren aus [3]. Die besten Werte sind **fett** markiert.
*Werte nicht genau bekannt.

Verfahren	Recall [%]	Precision [%]	F-Measure [%]	Laufzeit [s]
evolutionäres Graphmatching	**71,63**	**55,00**	**62,22**	**4,23**
Hopfieldmatcher NNGM aus [3]	70,50	46,26	55,58	20 - 50*
Similarity Flooding aus [3]	69,83	49,84	58,16	5,41

maximalen Recall sondern für die anderen Gütemerkmale optimiert werden, fällt die jeweilige Verbesserung noch stärker aus. Vor dem Hintergrund des schlechten Signal-Rausch-Verhältnisses von ca. 19 Handknochen: 2.300 Regionen = 0,008 wird der Erfolg noch deutlicher.

Insgesamt konnte also für die Objektidentifikation ein bislang schon in anderen Kontexten angewendetes Verfahren für die Verwendung mit Strukturprototypen angepasst werden. Die bisherigen Ergebnisse wurden in allen Gütekriterien deutlich verbessert obwohl gleichzeitig die Anzahl der relevanten Verfahrensparameter auf lediglich drei gesenkt werden konnte.

Für zukünftige Arbeiten gilt es die Segmentierungs- und damit auch die Trainingsqualität zu steigern. Zudem kann eine Einteilung in verschiedene geschlechtsabhängige Altersklassen sinnvoll sein, da sich die Ausprägung der Handknochen und damit die aus den Röntgenbildern extrahierten Textur- und Formmerkmale entsprechend stark unterscheiden [8]. Bislang wurde auch noch keinerlei Vorverarbeitung in Bezug auf Histogrammanpassung oder Bildausrichtung unternommen, wodurch weitere Verbesserungen zu erwarten sind. Letztendlich gilt es noch, neben der Identifikation von Handknochen weitere Anwendungsgebiete mit dem Verfahren zu testen.

Literaturverzeichnis

1. Metzen J, Krüger T, Schenk A, et al. Matching von Baumstrukturen. Proc BVM. 2007; p. 116–120.
2. Braumann UD, Franke H, Hengstler J, et al. Graph-based quantification of astrocytes. Proc BVM. 2006; p. 379–383.
3. Fischer B, Sauren M, Güld M, et al. Scene analysis with structural prototypes for content-based image retrieval in medicine. Procs SPIE. 2008;6914.
4. Conte D, Foggia P, Sansone C, et al. Thirty years of graph matching in pattern recognition. Intern J Pattern Recognit Artif Intell. 2004;18(3):265–298.
5. Pelillo M. Matching free trees, maximal cliques and monotone game dynamics. IEEE Trans PAMI. 2002;24(11):1535–1531.
6. Pavan M, Pelillo M. Generalizing the Motzkin-Straus theorem to edge-weighted graphs, with applications to image segmentation. LNCS. 2003; p. 485–500.
7. Bomze I. Evolution towards the maximum clique. J Glob Opt. 1997; p. 143–164.
8. Greulich W, Pyle S. Radiographic atlas of skeletal development of the hand and wrist. Palo Alto, CA, USA: Stanford University Press; 1959.