Kategorisierung von digitalen Röntgenbildern mit parametrisierbaren Formmodellen

Michael Kohnen, Frank Vogelsang, Frank Weiler, Jörg Bredno*
und Jörg Dahmen†

Klinik für Radiologische Diagnostik
*Institut für Medizinische Informatik
†Lehrstuhl für Informatik VI
Rheinisch-Westfälische Technische Hochschule (RWTH), 52057 Aachen
Email: kohnen@rad.rwth-aachen.de

Zusammenfassung. Die automatische Kategorisierung von Bildern im Zusammenhang mit digitalen Bildarchiven erlangt in der medizinischen Informatik eine immer größere Bedeutung. Wir verwenden Vorwissen über die möglichen dargestellten Objekte und deren spezifischen Formmerkmale. Das Verfahren ist in der Lage, unterschiedliche Formeigenschaften anhand einer Trainingsdatenmenge mit Hilfe von vergleichsweise wenigen Parametern hinreichend genau zu beschreiben. Eine Optimierung dieser Formmodelle liefert für jedes Modell eine Minimalenergie, die die Korrelation des Modells mit den abgebildeten Objekten im Bild beschreibt. Anhand dieser Formenergien kann eine Zuordnung des Bildes in eine bestimmte Kategorie erfolgen. Entgegen klassischen Ansätzen, die kein Vorwissen einsetzen, erweist sich dieser Ansatz als robust gegenüber Bildartefakten und unvollständiger Objektkonturinformation.

Schlüsselwörter: Image Retrieval, Kategorisierung, Formmodell, Simulated Annealing

1 Einleitung

In der medizinischen Informatik erlangen digitale Bildarchive eine immer größere Bedeutung. Das IRMA-Projekt (Image Retrieval in Medical Applications) hat unter anderem zum Ziel, Methoden zu liefern, die eine automatische Kategorisierung sekundär digitalen Bildmaterials ermöglichen. Mit Hilfe des *Active Shape Model* [1] wird eine Zuordnung der Bilder auf Grundlage von Objektkonturen unter Einbeziehung von Vorwissen ermöglicht. Die automatische Kategorisierung wird benötigt, um später unter anderem Archivanfragen diagnostischer Relevanz zu ermöglichen.

Parametrisierbare Formmodelle werden in der medizinischen Bildverarbeitung bisher zur Segmentierung des Bildmaterials herangezogen. Hier wird die Kategorisierung von Röntegenaufnahmen als ein alternatives Anwendungsgebiet parametrisierbarer Formmodelle aufgezeigt.

2 Methode

Bei dem hier verwendetem Active Shape Model handelt es sich um eine als geschlossenes Polygon dargestellte Kontur v, für die durch Paramteroptimierung die bestmögliche Übereinstimmung mit im Bild vorhandenen Konturen aufgefunden wird. Die Optimierung findet auf einem Bildpotential D statt.

Unter der Voraussetzung, daß eine Kategorisierung des Bildmaterials in eine fest vordefinierte Anzahl von Kategorien $K_i, i \in \{1, \ldots, N\}$ erfolgen soll, können Forminformationen über die in der jeweiligen Kategorie auftretenden Bildobjekte a-priori trainiert werden. Das Active Shape Model ist in der Lage, in Abhängigkeit von vergleichsweise wenigen Parametern mögliche Formänderungen, präferiert durch eine Menge von Trainingsformen, eines gesuchten Objektes zu modellieren. Dabei wird die große Variabilität biologischer Strukturen durch eine Trainingsdatenmenge reflektiert.

2.1 Berechnung der Formmodelle

Das Verfahren der Active Shape Models versucht durch Dimensionsreduktion des Merkmalsraumes, der durch die Trainingsdatenmenge gegeben ist, die charakteristischen Formparameter mittels einer Hauptachsentransformation zu bestimmen. Um ein Active Shape Model zu berechnen, wird eine affin normierte Stichprobenmenge, bestehend aus m Konturen mit jeweils n sortierten Stützpunkten, benötigt. Eine affine Normierung der Stichprobenmenge kann mit Hilfe der Fehlerquadratmethode bestimmt werden.

Eine Kontur v_i der normierten Stichprobenmenge kann als Vektor $\mathbf{v}_i$ der Länge $2n$ definiert werden, dessen Einträge die Koordinaten seiner Stützpunkte sind:

$$\mathbf{v}_i = (x_{i_1}, y_{i_1}, x_{i_2}, y_{i_2}, \ldots, x_{i_n}, y_{i_n}) \tag{1}$$

Die Konturen der Stichprobenmenge können als Elemente eines $2n$-dimensionalen Vektorraums aufgefaßt werden. Das Ziel ist nun, den durch die charakteristischen Formmerkmale aufgespannten Unterraum anhand weniger Basisvektoren zu beschreiben. Dazu wird die $2n \times 2n$-Kovarianzmatrix K generiert, deren Einträge k_{ij} die Abhängigkeiten der jeweiligen Konturpunkte zueinander modellieren:

$$k_{ij} = \frac{1}{2n} \sum_{l=1}^{2n} (v_{l_i} - \overline{v}_i)(v_{l_j} - \overline{v}_j), \tag{2}$$

wobei $\overline{v}_i$ die Mittelwertkontur der Stichprobenmenge darstellt.

Die Hauptachsentransformation der Kovarianzmatrix K liefert eine Orthonormalbasis von Eigenvektoren mit ihren jeweiligen Eigenwerten. Die Größe der einzelnen Eigenwerte entspricht der Varianz der Stichprobenmenge entlang des zugehörigen Eigenvektors. Die Eigenvektoren werden nach der Größe ihrer Eigenwerte sortiert.

Eine Formvariation v^* des Modells kann mit Hilfe der Orthogonalbasis P, bestehend aus den k Eigenvektoren mit den größten Eigenwerten, und den zugeordneten variierten Eigenwerten ξ_i durch

$$v^* = \overline{v} + P\xi = (\overline{v}_1 \ldots \overline{v}_n) + \left[\begin{pmatrix} p_{1,1} & \cdots & p_{1,k} \\ \vdots & & \vdots \\ p_{n,1} & \cdots & p_{n,k} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_k \end{pmatrix} \right]^T \tag{3}$$

berechnet werden.

Translation, Skalierung und Rotation wird durch eine auf v^* angewendete affine Transformation modelliert. Das so erhaltene Modell ist damit affin invariant und kann die gewünschten Formvariationen berücksichtigen.

In der Praxis nimmt die Größe der Eigenwerte schnell ab. Deshalb reichen wenige Eigenvektoren der Orthonormalbasis aus, um die auftretenden Formvariationen zu modellieren. Daher können die hier verwendeten Formmodelle durch zehn Parameter mit ausreichender Genauigkeit berechnet werden. Dies sind fünf affine Parameter (Translation in x- und y-Richtung, Skalierung in x- und y-Richtung, Rotation um den Winkel α) und die Variationen der fünf größten Eigenwerte. Die zulässige Variationsbreite der Parameter wird durch eine Festlegung von Grenzen eines diskretisierten Intervalls bestimmt.

2.2 Generierung des Bildpotentials

Zunächst ist eine Vorverarbeitung des Bildmaterials notwendig, um als Ausgangsbasis für eine Optimierung des Active Shape Model dienen zu können. Ziel dieser Vorverarbeitung ist die Generierung eines Bildpotentials, daß möglichst wenige lokale Minima enthält, um das Optimierungsverhalten des Formmodells zu verbessern.

Die im Bild enthaltene Konturinfomation wird mit Hilfe des Canny-Edge-Detectors [2] extrahiert. Da dieser Kantenfilter anhand von Grauwertgradienten die Zuordnung der Bildpixel zu Kanten vornimmt, wird eine Vielzahl von irrelevanten Kanten detektiert, die unter anderem durch Bildstörungen und andere Artefakte (z.B. medizinische Marker, Metallimplantate, etc.) hervorgerufen werden. Daher werden aus dieser primären Kantenmenge Kanten aussortiert, die bestimmten Kriterien, wie minimale Länge oder Inhomogenität des Kantenverlaufs, nicht genügen.

Als Bildpotential für die Optimierung wird schließlich ein distanztransformiertes Kantenbild verwendet. Die Grauwerte eines distanztransformierten Kantenbildes entsprechen der City-Block-Distanz zur nächsten Bildkante. Das Distanzbild hat den Vorteil, daß für jeden Pixel in $O(1)$ der Abstand zum nächsten lokalen Minimum bestimmt werden kann.

2.3 Optimierung der Formmodelle

Ziel der Optimierung ist die Ermittlung derjenigen Formvariation, die unter Berücksichtigung des Bildpotentials D eine minimale Energie E_{v^*} aufweist. Die

Energie der Kontur wird durch eine gewichtete Summation der von den Knoten des Polygons überdeckten Pixelgrauwerte bestimmt:

$$E_{v^*} = \sum_{i=1}^{n} w_i v_i^* \cdot D_{xy}, \quad \text{mit} \quad \sum_{i=1}^{n} w_i = 1 \tag{4}$$

Da die Anzahl der Variationen eines Active Shape Models mit k Parametern, wobei jeder Parameter l verschiedene Werte annehmen kann, gerade l^k ist, fällt eine exakte Lösung des Optimierungsproblems in die Klasse der NP-vollständigen Probleme.

Daher wird ein stochastisches Optimierungsverfahren gewählt, das sich in der Praxis bei der Lösung solcher Probleme oft bewährt hat, das Simulated Annealing. Voraussetzung für die Anwendbarkeit dieses Verfahrens ist jedoch eine topologische Ordnung des Suchraums. Die topologische Ordnung wird durch eine Nachbarschaftsbeziehung modelliert. Im Fall der hier verwendeten Formmodelle besitzt jede Variation genau zwanzig Nachbarkonfigurationen, da jeder der 10 Parameter durch Erhöhen bzw. Erniedrigen eine neue Konfiguration beschreibt. Die Suche erfolgt über eine zufallsbasierte Bestimmung neuer Nachbarn deren Akzeptanzwahrscheinlichkeit von einer Boltzmannverteilung abhängt.

Dadurch ist das Formmodell in der Lage, lokale Minima zu überwinden und unter sukzessiver Bestimmung neuer Nachbarkonfigurationen gegen das globale Minimum zu streben. Da das Bildpotential in der Regel eine Vielzahl von lokalen Minima enthält ist Simulated Annealing ein probates Verfahren zur Ermittlung des globalen Optimums.

2.4 Kategorisierung des Bildmaterials

Zur automatischen Kategorisierung eines Bildes wird für jede Klasse ein Active Shape Model erzeugt und mit Aufnahmen trainiert, die die biologische Variabilität in den einzelnen Klassen abbilden.

Nach Optimierung aller Formmodelle auf dem zu kategorisierenden Bild liegt für jedes Formmodell ein Energiemaß E_{v^*} vor, daß ein Maß für die Paßgenauigkeit des Formmodells auf dem Bild darstellt. Das Bild soll nun anhand der verschiedenen Energien einer bestimmten Kategorie zugeordnet werden. Da die Formmodelle jedoch unterschiedlich viele Stützpunkte besitzen, ist eine Normierung der Energien bezüglich der Punktanzahl notwendig, um eine Vergleichbarkeit zu erzielen.

Intention dieses Ansatzes ist, daß genau das Formmodell die geringste Energie aufweist, daß mit den auf dem Bild dargestellten Objekten am besten korreliert. Eine Zuordnung des Bildes kann damit anhand des minimalen Energiemaßes aller Formmodelle erfolgen.

3 Ergebnisse

Bisher wurden zwei Active Shape Models generiert, die jeweils die charakteristischen Formmerkmale von Händen [3] und von Wirbelkörpern modellieren. Die

Variabilitäten wurden durch ein Training mit jeweils 12 Aufnahmen bei beiden Modellen analysiert. Die Anwendung nicht trainierter Aufnahmen aus einer der beiden Klassen ergab beim Optimieren des Hand-Modells auf 20 verschiedenen Hand-Aufnahmen ein mittleres Energiemaß von $\mu_h = 0,280$ mit einer Standardabweichung von $\sigma_h = 0,169$. Wenn dagegen das Wirbelmodell auf den gleichen Handaufnahmen angewendet wurde erhöhte sich die mittlere Restenergie auf $\mu_w = 0,556$ mit einer Standardabweichung von $\sigma_w = 0,258$. Dies zeigt die Leistungsfähigkeit des Ansatzes bei der Unterscheidung unterschiedlicher Bildkategorien.

Desweiteren ist das Verfahren relativ unempfindlich gegenüber Bildartefakten, wie medizinischen Markern oder metallischen Implantaten. Falls jedoch zu große Teile der relevanten Konturen überlagert sind, oder das Objekt atypische Formmerkmale aufweist, die nicht in der Trainingsdatenmenge enthalten sind, kann es nicht zuverlässig detektiert werden und damit keine korrekte Zuordnung stattfinden.

4 Diskussion

Dieser Ansatz versucht durch Einsatz von a-priori Wissen über Objektkonturen auf dem Bildmaterial eine Kategorisierung des Bildmaterials zu erreichen. Die Formbasierte Kategorisierung ohne Verwendung von Formwissen führt bisher nicht zu befriedigenden Ergebnissen [4]. Jedoch ist anzumerken, daß bei Bildern, die keine oder zu verrauschte Kanteninformation enthalten (z.B. Ultraschallaufnahmen), das Verfahren nur bedingt anwendbar ist. Ebenfalls treten Probleme auf, wenn die Variationsbreite der dargestellten Objekte zu groß ist, wie zum Beispiel bei Kontrastmittelaufnahmen des Darms. Daher sieht das IRMA-Konzept zur Klassifikation eine Kombination verschiedener Methoden vor, um eine Verbesserung der Klassifikationsergebnisse zu erzielen.

Nichtsdestoweniger zeigt das Verfahren interessante Möglichkeiten auf, Formwissen, in einen Kategorisierungsprozeß zu integrieren. Insbesondere die Unempfindlichkeit des Verfahrens gegenüber Bildartefakten, hebt es von klassischen Verfahren ab. Darüber hinaus liefert das Verfahren eine grobe Vorsegmentierung der dargestellten Objekte, die sich für nachfolgende Segmentierungsverfahren als hilfreich erweisen kann.

Literatur

1. Cootes T, Taylor C, Cooper D, Graham J: Active Shape Models - Their Training and Application. Computer Vision and Image Understanding Vol. 61, No. 1, Jan, pp. 38-59, 1995
2. Canny J: A computational Approach to Edge Detection. IEEE Transactions PAMI, Vol 8(6), p. 679, 1986
3. Kohnen M: Modell- und wissensbasierte Segmentierung von Handröntgenbildern. Diplomarbeit Naturwissenschaftliche Fakultät, RWTH Aachen, 1998
4. Bredno J, Brandt S, Dahmen J, Wein B, Lehmann T: Kategorisierung von Roentgenbildern mit aktiven Konturmodellen. (Dieser Band).