



Management und Verarbeitung medizinischer multimedialer Daten

Henning Müller, Thomas Deselaers, Thomas M. Lehmann, Antoine Geissbühler

Zusammenfassung

Dieser Artikel gibt eine Einführung in die inhaltsbasierte Suche in medizinischen Bilddatenbanken. Es wird klar, dass neue bildgebende Verfahren in der Medizin nicht nur für eine erhöhte Variabilität der Bild- oder Multimedia-Daten sorgen, sondern auch die Anzahl der erzeugten Dokumente und deren Größe ständig steigt (Radiologie Genf: ~40.000 Bilder pro Tag im Jahr 2006, [20]). Wie in der Radiologie, wo mittlerweile fast alle Bilder in digitaler Form verfügbar sein können, werden auch in anderen Bereichen (Pathologie, Dermatologie) verstärkt Bilder digital erzeugt, die über kurz oder lang ebenfalls im PACS (Picture Archival and Communication System) gespeichert werden müssen.

Bisher wurden medizinische multimediale Daten nur patientenbezogen gespeichert und genutzt. Da in den Daten vergangener Fälle auch viel Information für die Zukunft gespeichert ist, drängt die Forschung verstärkt dazu, Daten gezielt für die Diagnoseunterstützung zu extrahieren und zu benutzen. Das birgt auf der einen Seite Chancen für eine verbesserte Diagnoseunterstützung, auf der anderen Seite Risiken für die Privatsphäre der Patienten. Bei der Erstellung von Datenbanken muss also vorsichtig vorgegangen werden, und vor allem müssen an Kliniken klare Richtlinien für die Weiterbenutzung von klinischen Daten existieren um die geplante Benutzung zu erleichtern. Inhaltsbasierte Bildsuche ist eine Technik, die es erlaubt, in Datenbanken nach visuell ähnlichen Bildern zu suchen, bzw. im medizinischen Bereich Fälle mit ähnlichen Bildern zu finden [11, 16]. Dies entspricht einem Mediziner, der alte Fälle ähnlich zu seinem aktuell behandelten Fall benötigt, um sie mit der jetzigen Situation zu vergleichen. Alte Fälle müssen dabei anonymisiert sein, allerdings so viel Information wie möglich

enthalten, z. B. Bilder, Alter, Medikamente, Diagnose, Behandlung aber auch, falls möglich, das Resultat der Behandlung.

Dieser Artikel führt multimediale Datenquellen im Krankenhaus ein und beschreibt die inhaltsbasierte Suche mit all ihren Vorteilen und Problemen. Am Ende werden Forschungsrichtungen beschrieben, die den aktuellen Techniken weiterhelfen dürften.

1 Einführung

Multimediale Daten (Bilder, EKG-Kurven, Videos, aber auch Texte und strukturierte Daten) spielen in der medizinischen Diagnostik und Behandlung von Patienten eine sehr wichtige Rolle. Da diese Daten mittlerweile überwiegend in elektronischer Form abgelegt werden, werden sie auch für eine weitergehende Analyse verfügbar, anders als es vorher bei der papierbasierten Akte der Fall war. Bisher sind solche Analysen vor allem mit strukturierten Daten zur Diagnoseunterstützung gemacht worden. Die Analyse von Freitext ist deutlich schwieriger, aber ebenfalls gut möglich und wird häufig angewandt. Visuelle Analyse medizinischer Multimediadaten wird bisher überwiegend pro Patient gemacht, allerdings selten über voluminöse Datenbanken mit einer großen Anzahl an Patienten. Dies ändert sich zur Zeit. Evidence-based medicine [1] oder case-based reasoning sind zwei Techniken, die bisher überwiegend mit strukturierten Daten oder Freitext arbeiten, aber in einige Fällen auch schon mit visuellen Daten möglich sind, oder zumindest visuelle Daten enthalten [2].

Bild „information retrieval“ ist ursprünglich vor allem für photographische Datenbanken angewandt worden [11], in denen Journalisten z. B. nach Bildern für die Illustration eines Artikels suchen. Während die Suchmaschinen im Web, wie z. B. Google und Alltheweb, sich vor allem auf den Text neben den Bildern konzentrieren, konzentrieren sich visuelle Suchmaschinen auf den Inhalt der Bilder selbst, also die Verteilung von Farben im Bild oder Formen von Objekten. Dies ermöglicht eine andere Form der Navigation anhand von Beispielbildern („Zeige mir Bilder so ähnlich wie dieses aber anders als jenes“). Ein anderer Anwendungsbereich für eine solche Ähnlichkeitssuche ist die Suche nach Markenzeichen. Während Objekte in Bildern häufig eine wichtige Rolle spielen sind viele Markenzeichen sehr abstrakt und schwierig mit Worten zu beschreiben. Visuelle Suche kann in diesen Fällen helfen.

Im medizinischen Umfeld wurde eine solche visuelle Navigation bereits häufig gefordert, allerdings nur recht selten wirklich in die Praxis umgesetzt. Erste Anwendungsbereiche waren die Klassifikation von Bildern, z. B. um Fehler in den DICOM-Headern (Digital Imaging and Communications in Medicine) zu finden und automatisch zu korrigieren. Andere Einsatzbereiche sind eine Diagnoseunterstützung in eng begrenzten Bereichen (Lungen CTs [2], Dermatologie [25], Pathologie etc.), in denen Datenbanken speziell für diesen Zweck erstellt werden.

Mittlerweile scheinen die Techniken reif für einen Einsatz in größeren Datenbanken, die mittlerweile mehr oder

Autoren: Henning Müller, Thomas Deselaers, Thomas M. Lehmann, Antoine Geissbühler

Titel: Management und Verarbeitung medizinischer multimedialer Daten

In: Jäckel (Hrsg.) Telemedizinführer Deutschland, Ober-Mörlen, Ausgabe 2007

Seite:



Bildgestützte Diagnostik, Teleradiologie

weniger automatisch in vielen Kliniken entstehen, um Forschung und Lehre zu unterstützen [16]. Diese Datenbanken können entweder getrennt vom Bildarchiv entstehen, oder aber eine direkte Verbindungen zum Bildarchiv haben. Wichtig ist es dabei vor allem von Seiten der Institution diese Anstrengungen zu unterstützen und eine Infrastruktur zum anonymisierten Speichern der Bilder zu Verfügung zu stellen, um die Entstehung eigener, privater Datenbanken zu verhindern und stattdessen eine übergreifende Datenbank in der Institution zu erstellen. Dies erlaubt es auch, die Privatsphäre der Patienten besser zu schützen, indem generelle Tools zur Anonymisierung der Daten bereitgestellt werden. Außerdem kann eine Prozedur zur Veröffentlichung der Daten die Kontrolle und ebenfalls die Datenqualität erhöhen. Im selben Moment kann eine solche Infrastruktur die Breite der multimedialen Daten erfassen, wie es in Abschnitt 2.1 versucht wird.

Generell sollte dieser Artikel klar machen, dass multimediale medizinische Daten nicht separat betrachtet werden können, sondern im Zusammenhang mit allen anderen Daten über den Patienten. Suchmethoden für medizinische Fälle müssen Daten verbinden und in diesen Daten die relevanten Punkte finden. Dies gelingt nicht nur durch eine visuelle Analyse sondern durch ein Verbinden komplexer Datenquellen.

2 Multimediale Datenquellen und Datenbanken

Es ist schwierig alle multimedialen Daten die im Krankenhaus anfallen zu erfassen und so kann die folgende Liste hier auch nur einen Teil der Daten wiedergeben, um die Komplexität des Problems zu erläutern und bei der Erstellung von Infrastrukturen zu helfen. Dieses Kapitel bezieht sich überwiegend auf Erfahrungen aus der Radiologie des Genfer Universitätsspitals mit dem Versuch, einige Aussagen zu verallgemeinern.

2.1 Datenquellen

Der Begriff Multimedia kann generell als die Verbindung von mehreren Medien verstanden werden (Text, Bild, Ton, etc.). Generell wird der Begriff aber häufig in ei-

ner abgeschwächten Form benutzt und so wird ein Bild (ein Medium) auch gerne als Multimedia bezeichnet. Man kann sagen, dass die Patientenakte eine multimediale Akte ist, die eine Vielzahl von Daten beherbergt, die in immer größerem Maße digital verfügbar sind. Die untenstehende Liste versucht, einen Teil dieser Daten nach Komplexität zu ordnen, um die Probleme besser verstehen zu können. Diese Liste ist nicht vollständig und dient daher eher als Startpunkt. Jede Institution wird sicherlich die eine oder andere Ähnlichkeit aber auch Unterschiede haben.

- Strukturierte Daten: Laborwerte, Patientendaten, administrative Daten usw.
- Freitext: Entlassungsbrief, Anamnese, Radiologiebrief, usw. (häufig auch im PDF-Format oder ähnlichem)
- Eindimensionale Signale über die Zeit: EKG, EEG, usw.
- Zweidimensionale Signale (Bilder): Röntgen, Ultraschall, Pathologiebilder (mikroskopisch), Dermatologie (Melanombilder, Fotos), usw.
- Zusammengehörige Bildserien: Studien mit Ganzkörperbildern, usw., z. B. in der Dermatologie in Genf zur Melanombeobachtung mit 30 Bildern desselben Patienten, um den gesamten Körper abzudecken.
- 3-dimensionale Schichtbilder: Daten aus CT, MR, PET, SPECT, usw.
- Videos: Schlaflabor, Endoskopie, Kardiologie, aber auch Ultraschall, usw.
- 3-dimensionale Bildansichten: Ansichten und Rekonstruktionen von Volumendaten oder Ultraschall, z. B. Oberflächenbilder, usw.
- 4-dimensionale Bilder: Flussimulationen (z. B. Aneurismen) über dreidimensionale Bilder, zeitgetriggerte Tomographien um die Herzbewegung darzustellen, usw.
- N-dimensionale Kombinationen: PET/CT, kombinierte Modalitäten, die zusammen registriert werden.

Diese Liste kann natürlich nicht genau auf die Einzelheiten der Bilder eingehen, sondern gibt eher einen groben Überblick. Einige Modalitäten erscheinen überraschend, wie die Videos, die das Schlaflabor in Genf produziert. Jede Nacht werden die Bewegungen von fünf Patienten über die

ganze Nacht digital aufgezeichnet und dann in Auszügen auf DVDs gespeichert. Insgesamt werden so pro Jahr weit über 1.000 DVDs produziert. Jede Institution hat so ihre eigenen Datenquellen, die verarbeitet werden müssen.

Ein anderer Bereich ist auch die Tendenz in Kliniken, den Patienten Fernsehen über Netzwerk zur Verfügung zu stellen. Solche Streaming-Applikationen sind zwar nicht für die Behandlung des Patienten nötig, werden aber eventuell mit derselben Infrastruktur realisiert und müssen entsprechend integriert werden.

2.2 Datenbanken

Die optimale Lösung ist natürlich das Speichern aller vorhandenen Daten in einer einzigen Infrastruktur, dem PACS (Picture Archival and Communication System). In Genf ist das Ziel eine entsprechend uniforme Infrastruktur. Ein Problem sind dabei die schieren Datenmengen. Im Moment produziert allein die Radiologie etwa 40.000 Bilder pro Tag, was zu einem jährlichen Volumen von mehreren TeraByte führt. Alle Daten müssen mindestens 10 Jahre lang zugänglich sein. Performance-Probleme treten bereits heute auf, und bisher ist neben der Radiologie nur die Kardiologie direkt ans PACS angeschlossen. Weitere Abteilungen sollen folgen und werden das Datenvolumen ebenso erhöhen wie neue bildgebende Modalitäten.

Neben dem PACS gibt es in Genf noch eine weitere Struktur zur Speicherung von Fällen vor allem für Forschung und Lehre. Die Datenbank heißt Casimage¹, ist von einem Radiologen entwickelt worden und ermöglicht das Speichern von Fällen mit textueller Annotation (siehe Abbildung 1) [24]. Bisher sind hier über 70.000 Bilder intern im Klinikum abgelegt, und per Internet ist eine Auswahl von 9.000 anonymisierten Bildern frei verfügbar.

Bei einer Umfrage zur Benutzung von Bildern im Uniklinikum kam heraus dass sowohl Casimage also auch das PACS sehr häufig benutzt werden, dass es aber auch noch viele andere Arten der Bildspeicherung gibt. So betreiben mehrere Abteilungen noch immer eine Speicherung auf externen Medien, wie z. B. das Schlaflabor oder die Kardiologie, die Videos auf DVD speichern.

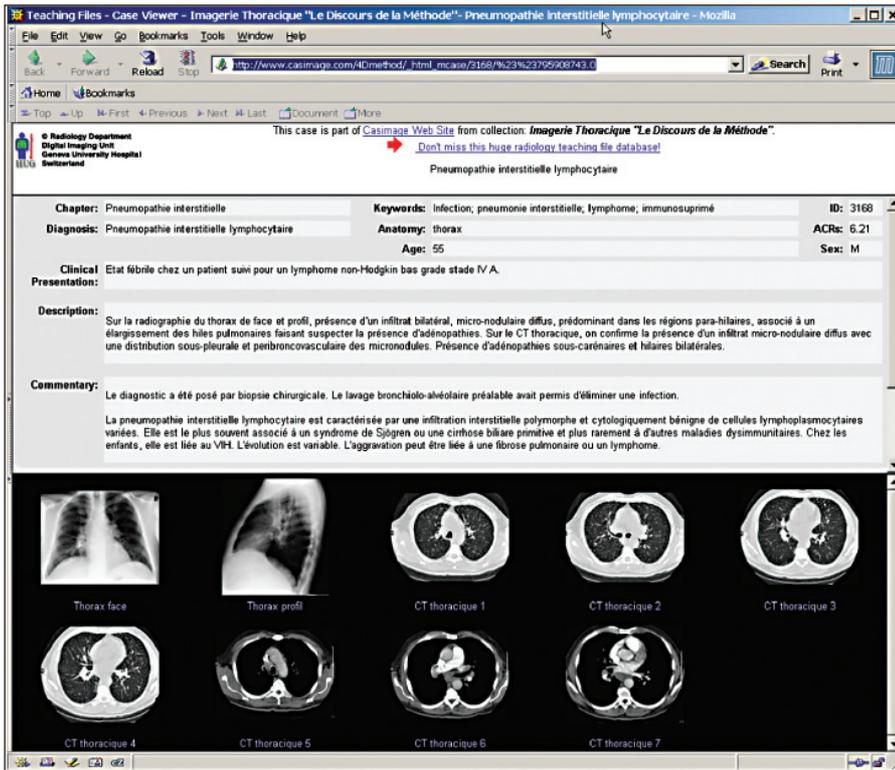


Abbildung 1: Das Interface der Casimage Bilddatenbank zur Speicherung von Fällen für die Lehre.

Die Pathologie unterhält zusätzlich eine eigene Bilddatenbank, von der aus Bilder in die Berichte kopiert werden können, und die auch von der Patientenakte aus zugänglich sind. Ein Anschluss an das PACS ist gewünscht, ist aber im Moment noch nicht möglich. Ähnlich verhält es sich in der Dermatologie, wo durch digitale Kameras die Hautbilder mittlerweile digital erzeugt werden.

Für die Forschung und Lehre werden neben der offiziellen Datenbank auch immer noch Verzeichnisse auf privaten Rechnern benutzt, da dies als einfacher für die Speicherung angesehen wird. Teilweise werden auch einfach die Namen der Patienten mit Charakteristiken des Falles notiert und die Daten dann bei Bedarf direkt aus dem PACS bezogen. Solche privaten Datenbanken sind natürlich nicht für die Allgemeinheit verfügbar und führen zu einer Aufsplitterung des Wissens.

Die Idee eines großen Wissensarchivs mit Zugang zu multimedialen Daten inklusive Text wäre natürlich die optimale Lösung. Wissen über Fälle von der Diagnostik, über die Behandlung bis hin zum Resultat der Behandlung wären optimal und könnten

sowohl in Forschung also auch bei der Lehre wichtige Resultate liefern.

3 Inhaltbasierte Suche

Normalerweise ist das Szenario der inhaltsbasierten Suche die Suche nach einem ähnlichen Beispielbild (Query-by-example) [26]. Abbildung 2 zeigt ein entsprechendes Szenario, in dem ein Me-

diziner einen Fall zu behandeln hat, dessen Diagnose er finden möchte. Er wählt ein Bild des Patienten aus und gibt es an die visuelle Suche, die ihm die ähnlichsten gespeicherten Fälle aus der internen Datenbank und deren Diagnosen zurückliefert. Auch wenn nicht unbedingt alle gefundenen Fälle dieselbe Diagnose haben, kann der Mediziner doch die gefundenen Fälle mit seinem aktuellen Fall vergleichen und so eventuell wichtige Informationen über die Diagnose und die Behandlung finden. Visuelle Suche in medizinischen Datenbanken ist häufig gefordert, aber leider nur selten umgesetzt worden [17, 22, 28].

Bisher wird in vielen bildlastigen Bereichen eines Krankenhauses sehr ähnlich gesucht. In der Radiologie oder der Pathologie ist es durchaus üblich, in Büchern nach möglicherweise visuell ähnlichen Bildern zu suchen, nur hier ist man auf die Suche per Pathologie angewiesen, die für einen neuen Fall ja gerade gesucht wird.

3.1 Anwendungen

Der erste Anwendungsbereich für die inhaltsbasierte Suche wird sicherlich in der Lehre zu finden sein, da hier die Applikationen weniger kritisch sind. Selbst einige falsch positive Bilder in den Resultaten machen ähnlich wie bei einer Textsuche in Google keine größeren Probleme. Man muss einfach nur die relevanten gefundenen Daten genauer betrachten. Für die Lehre kann es außerdem sehr interessant sein, visuell ähnliche Bilder mit einer unterschiedlichen Diagnose zu finden, um so die wichtigen Punkte besser verdeutlichen

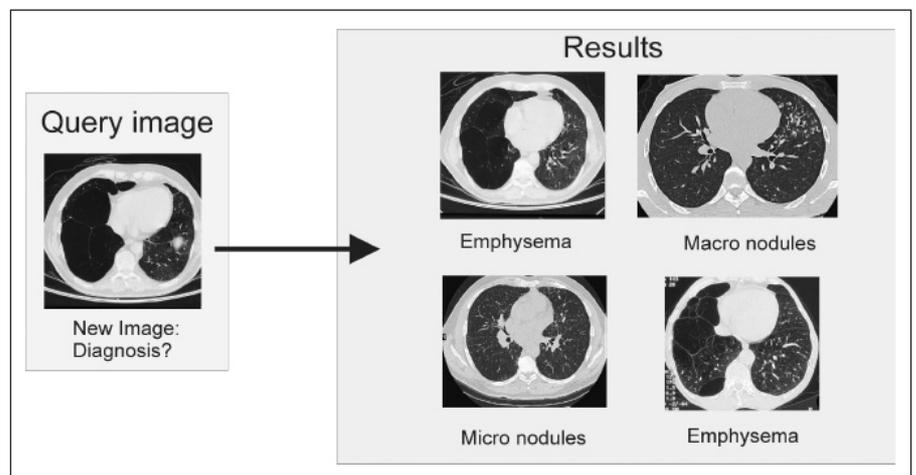


Abbildung 2: Ein Szenario für die Suche nach ähnlichen Fällen mit einem Beispielbild.

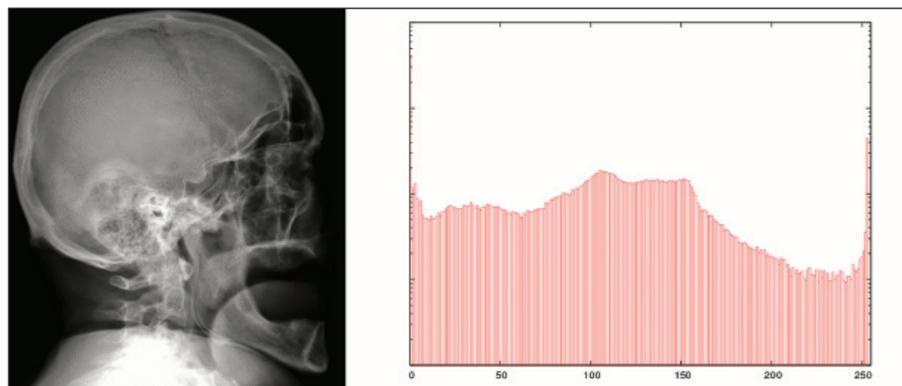


Abbildung 3: Ein Röntgenbild und das entsprechende Grauwerthistogramm. Man sieht, dass im Bild viele schwarze Pixel vorkommen (ganz links im Histogramm).

zu können. Voraussetzung für eine solche Suche sind natürlich verfügbare Datenbanken. Die RSNA² (Radiological Society of North America) versucht mit dem MIRC³ (Medical Image Resource Center) eine Infrastruktur zu schaffen, die es global ermöglicht Fälle mit Bildern verfügbar zu machen. Andere Initiativen in diesem Bereich sind auch HEAL⁴ (Health Education Assets Library) oder PathoPic⁵.

Für die Forschung kann die visuelle Suche ebenfalls interessant sein. Zum einen können Fälle für bestimmte Studien auch basierend auf visuellen Merkmalen ausgewählt werden, zum anderen können visuelle Merkmale direkt ausgewertet werden, ähnlich wie strukturierte Daten. Dies kann auch zu einem visuellen data mining führen, mit dem unbekannte Zusammenhänge zwischen visuellen Daten und Diagnosen oder strukturierten Daten gefunden werden können.

Die Diagnoseunterstützung ist mit Sicherheit das langfristige und schwierigste Ziel der visuellen Suche, aber auch der Bereich mit dem potentiell größten Einfluss. Hierfür ist es wichtig, spezialisierte Datenbanken zu erstellen, die der Praxis entsprechen und möglichst viele Fälle enthalten. Danach muss soviel Wissen wie möglich von den medizinischen Experten transferiert werden, um die visuellen Charakteristiken für eine spezielle Anwendung zu optimieren und die Anzahl der falsch positiven Fälle zu minimieren. Typische Anwendungsgebiete sind die Pathologie, die Dermatologie, die Radiologie mit z. B. Lungen-CTs, aber auch die Chirurgie, in der Fälle mit ähnlichen Brüchen für die

Vorbereitung von Operationen gesucht werden können.

Eine letzte Anwendung sind Routineaufgaben, für die die Klassifizierung von Bildern in bestimmte Klassen benutzt werden kann. Dies kann z. B. eine automatische Kontrolle von DICOM-Headern in Bildern sein, da gerade das Feld Anatomie häufig Fehler in den DICOM-Daten aufweist.

Eine Übersicht von mehr als 200 Artikeln mit sehr unterschiedlichen Anwendungen findet sich in [19].

3.2 Techniken

Zur inhaltsbasierten Bildsuche gibt es in der aktuellen Forschung zwei grundlegende Ansätze: Der eine Ansatz verwendet eine Methode, die als Suche nach den nächsten Nachbarn bekannt ist. Dazu wird das Beispielbild mit jedem Bild in der Datenbank verglichen und es werden die Bilder als Ergebnis zurückgeliefert, die dem Anfragebild am ähnlichsten sind. Dazu ist es notwendig, das Bild durch geeignete Deskriptoren so darzustellen, dass ein Ähnlichkeitsvergleich möglich ist. Einige exemplarische Beispiele hierzu sind weiter unten zu finden. Der andere Ansatz zur inhaltsbasierten Bildsuche ist an den Verfahren orientiert, die auch in der Textsuche, wie z. B. bei den großen Internetsuchmaschinen, verwendet werden. Auch hierzu wird ein Bild durch geeignete Deskriptoren dargestellt, jedoch werden nicht zu jedem Bild alle Deskriptoren abgespeichert, sondern es wird zu jedem Deskriptor abgespeichert, in welchen Bildern er vorkommt. Wird nun

eine Suchanfrage gestellt, werden aus dem Beispielbild die Deskriptoren extrahiert und in den entsprechenden Dateien nachgeschaut, in welchen Bildern sich die gleichen Charakteristika wiederfinden. Auf diese Weise wird es möglich, sehr große Merkmalsräume effizient zu handhaben. Der Nächste-Nachbar-Ansatz wird z. B. vom IRMA-System [16] verwendet. Der von der Textsuche inspirierte Ansatz wird im GIFT/medGIFT-System [27] verfolgt.

Wie bereits oben erwähnt, ist es in beiden Ansätzen notwendig, die Bilder durch geeignete Deskriptoren für den Computer vergleichbar zu machen. Allgemein lassen Bilder sich über ihre Farben, Texturen, und Formen beschreiben. Unter Texturen versteht man die Beschaffenheit und die Struktur von Bereichen. Als Beispiel sei hier der Unterschied zwischen dem Fell eines Dalmatiners und eines Zebras genannt: Obwohl in beiden Fällen, der Weiß- und der Schwarzanteil etwa gleich sind, ist es für Menschen sehr einfach diese beiden Texturen zu unterscheiden.

Im Folgenden werden kurz einige Möglichkeiten zur Repräsentation von Bildern für die Bildsuche erläutert. Jede dieser Möglichkeiten lässt sich in beiden Ansätzen zur Bildsuche direkt verwenden. Eine Übersicht und ein Vergleich verschiedener Bildrepräsentationen findet sich in [8].

Farbdeskriptoren

Ein Histogramm ist eine Darstellung der Häufigkeitsverteilung von Werten. Ein Farbhistogramm ist demnach eine Darstellung der Verteilung welche Farben zu welchem Anteil in einem Bild vorkommen. Bei einem Grauwertbild, erlaubt das Histogramm eine Aussage darüber zu treffen, wieviele Pixel eines Bildes hell bzw. dunkel sind. In Abbildung 3 ist ein Bild und daneben das entsprechende Grauwert histogramm abgebildet. Auf die gleiche Art kann man auch die Verteilung von Farben in einem Bild darstellen. Da jedoch Farben üblicherweise über mindestens 3 Komponenten (z. B. Rotanteil, Grünanteil, Blauanteil) dargestellt werden, braucht man in diesem Fall ein dreidimensionales Histogramm.

Ein Nachteil von Histogrammen ist, dass die Information über die räumliche Anordnung der Farben im Bild vollständig verloren geht. Daher ließen sich durch ein



Abbildung 4: Beispielbilder für unterschiedliche Textureigenschaften: „links“ grobe Textur/feine Textur, „mittig“ hoher Kontrast, niedriger Kontrast, „rechts“ gerichtet, nicht gerichtet.

Farbhistogramm die oben bereits genannten Bilder eines Dalmatiners und eines Zebras nicht unterscheiden.

Eine Möglichkeit, die räumliche Information zumindest teilweise zu erhalten, ist nicht nur ein Histogramm für ein Bild zu erstellen, sondern das Bild vorher in z. B. 4 Teilbilder (oben/unten, links/rechts) zu zerlegen und dann für jedes dieser Teilbilder ein eigenes Histogramm zu berechnen.

Texturdeskriptoren

Als Textur bezeichnet man im Allgemeinen die Strukturierung eines Bildes oder einer Bildregion. Es gibt viele verschiedene Ansätze, Texturen im Computer zu beschreiben. Eine der am häufigsten verwendeten Darstellungen geht auf [29] zurück. Dabei wird versucht mit mathematischen Formulierungen die Eigenschaften einer Textur so zu beschreiben, wie auch Menschen sie empfinden. Es werden sechs Eigenschaften beschrieben: Rauheit (coarseness), Kontrast (contrast), Gerichtetheit (directionality), Linienartigkeit (line-likeness), Regelmäßigkeit (regularity), und Grobheit (roughness). In Experimenten stellte sich heraus, dass insbesondere die ersten drei, also Rauheit, Kontrast und Gerichtetheit für die Texturbeschreibung ausgesprochen wichtig sind. Abbildung 4 zeigt jeweils ein Paar von Beispielbildern für diese Textureigenschaften.

Wenn die interessierenden Textureigenschaften festgelegt wurden, ist z. B. eine Möglichkeit, diese für die Umgebung eines jeden Bildpunktes zu berechnen und dann ein Histogramm über diese Eigenschaften zu erstellen.

Andere Texturbeschreibungsmöglichkeiten verwenden Gabor Filter [21], Cooccurrence Matrizen [13] oder fraktale Dimension [23] zur Beschreibung von medizinischen und nicht-medizinischen Bildern.

Formdeskriptoren

Um die Formen der Objekte in einem Bild zu beschreiben, ist es notwendig, das Bild zuerst zu segmentieren, d. h. zu bestimmen an welcher Stelle welches Objekt ist und wo Bildhintergrund ist. Dies ist jedoch kein gelöstes Problem, denn die zuverlässige Segmentierung allgemeiner Bilddaten ist gleichzusetzen mit der Erkennung aller abgebildeten Objekte und damit wäre das Bildsuchproblem ja bereits gelöst. Nichtsdestotrotz gibt es zahlreiche Arbeiten, die sich mit der Beschreibung von Formen in Bildern befassen, wie z. B. in MPEG-7 [4]. Die meisten der Ansätze gehen davon aus, dass die Objektform bereits bekannt ist und extrahieren dann Merkmale aus dem Umriss, wie z. B. den maximalen Durchmesser und Fourierkoeffizienten für die Krümmungen.

Ein anderer Ansatz wird in [5] vorgestellt. Hierbei wird jedes Bild vollautomatisch in Regionen unterteilt die teilweise auch den Objekten in den Bildern entsprechen. Um nun Bilder zu vergleichen werden die Regionen der beiden Bilder paarweise verglichen und die Ähnlichkeiten geeignet kombiniert. Ein Bildsuchsystem, welches auf diesen Prinzipien beruht ist Blobworld⁶.

Abhängig von der gewünschten Anwendung eignen sich unter Umständen noch andere Deskriptoren zum Bildvergleich. Zur Suche nach Röntgenbildern eignet sich z. B. ein direkter Vergleich der Bilddaten ausgesprochen gut [15]. Die einfachste Möglichkeit hierzu ist, zwei Bilder auf gleiche Größe zu skalieren und dann pixelweise zu vergleichen. Es hat sich jedoch gezeigt, dass es von Vorteil ist, gewisse Verformungen innerhalb der Bilder zuzulassen und daher werden in [15] die Pixel eines Bildes mit allen Pixeln aus der direkten Umgebung in dem anderen Bild verglichen.

Eine anderer, moderner Ansatz, beruht auf der Annahme, dass Objekte in Bildern

aus Teilen bestehen, die mehr oder minder unabhängig modelliert werden können [9, 10, 12, 18]. Davon ausgehend, werden aus den Bildern Teilbilder extrahiert, die dann unabhängig voneinander modelliert werden um am Ende eine globale Entscheidung über die Ähnlichkeit zweier Bilder treffen zu können, bzw. um zu entscheiden, welche Objekte im Bild zu sehen sind.

Benutzerinterfaces

In Abbildung 5 ist ein typisches Benutzerinterface einer visuellen Bildsuchmaschine zu sehen. Die meisten Systeme bieten webbasierte Interfaces an, um die Suche in die verschiedensten Umgebungen zu integrieren.

Häufig ist es dabei möglich, externe Bilder als Beispiele anzugeben, und sich dann mittels relevance feedback (Markieren von Resultaten als relevant oder irrelevant) in der Datenbank zu bewegen, um weitere relevante Resultate zu finden und die Suche einzuzugrenzen.

Häufig sind dabei die Suchmaschine und das Interface voneinander getrennt und kommunizieren über XML-basierte Sprachen wie MRML⁷ (Multimedia Retrieval Markup Language).

3.3 Evaluierung

Ein wichtiger Punkt in Forschung und Entwicklung sind objektive Evaluationen von Systemen auf großen und realistischen Datenmengen, um es zu ermöglichen Systeme zu vergleichen und dadurch festzustellen, welche Ansätze vielversprechend sind und um neue Systeme mit dem Stand der Forschung vergleichen zu können und dadurch festzustellen, ob das neue System wirklich besser ist als das Alte. Im Bereich der natürlichen Sprachverarbeitung im Allgemeinen und bei der Textsuche im Speziellen existieren solche Strukturen schon länger, vor allem mit TREC⁸ (Text Retrieval Conference) und auch CLEF⁹ (Cross Language Evaluation Forum).

Durch Evaluationen wird die Forschung beschleunigt, denn es ist notwendig, die Ziele und Probleme klar zu definieren und auch die fehlenden Stücke eines Gesamtsystems zu identifizieren, um in Zukunft auf solche Lösungen hinarbeiten zu können.

Seit dem Jahr 2003 gibt als Teil der CLEF-Kampagne eine ImageCLEF¹⁰-



Bildgestützte Diagnostik, Teleradiologie

Initiative, deren Ziel es ist, einen Standardtestfall für inhalts- und textbasierte Bildsuche zu schaffen. Im Jahr 2003 war daher erstmals eine Evaluation vorgesehen in der das Ziel war, Bilder zu finden. Dazu wurden ca. 30.000 historische, schottische Bilder mit annotierendem Text zur Verfügung gestellt und zu bearbeitende Anfragen wurden in mehreren Sprachen definiert. Die Reaktionen auf diese Initiative waren sehr positiv, so dass im Jahr 2004 eine weitere Aufgabe zur visuellen und textuellen Suche nach medizinischen Bildern definiert wurde. Hierbei war es das erste Mal wirklich notwendig auch Bildverarbeitungsverfahren einzusetzen, da die Anfragen nur als Bilder vorhanden waren. Als Datenbank diente dazu der öffentliche Teil der bereits oben erwähnten Casimage-Datenbank und 26 Anfragebilder waren zu bearbeiten [7]. Ein Beispiel-fall aus der Casimage Datenbank ist in Abbildung 1 zu sehen.

Im Jahr 2005 wurde die medizinische Retrievalaufgabe mit einer Datenbank von über 50.000 Bildern durchgeführt und die Anfragen wurden sowohl durch Beispielbilder als auch durch beschreibenden Text definiert, so dass eine Kombination verschiedener Retrievaltechniken notwendig wurde. Außerdem wurde eine weitere medizinische Aufgabe eingeführt, bei der es um die Klassifikation bzw. Annotierung von medizinischen Daten ging [6]. Die weiter steigende Teilnehmerzahl in allen Aufgaben zeigt, wie wichtig solche Vergleichsaufgaben für die Forschungsgemeinschaft sind.

Im Jahr 2006 wurden die medizinischen Aufgaben ähnlich fortgesetzt wie in 2005, wobei jeweils neue Anfragebilder bzw. Informationsbedürfnisse verwendet wurden. Insbesondere an den Ergebnissen der Klassifikations-/Annotierungsaufgabe sieht man, welche großen Fortschritte im Bereich der Klassifikation medizinischer Röntgenbilder innerhalb eines Jahres möglich sind. Das System, das 2005 die besten Ergebnisse erzielte, ist 2006 nur noch im oberen Mittelfeld der Ergebnisse angesiedelt und neue Verfahren in denen das automatische Lernen im Vordergrund steht und die auf der oben erwähnten Annahme, dass Objekte in Bildern aus Teilen bestehen beruhen, konnten deutlich bessere Ergebnisse erzielen.

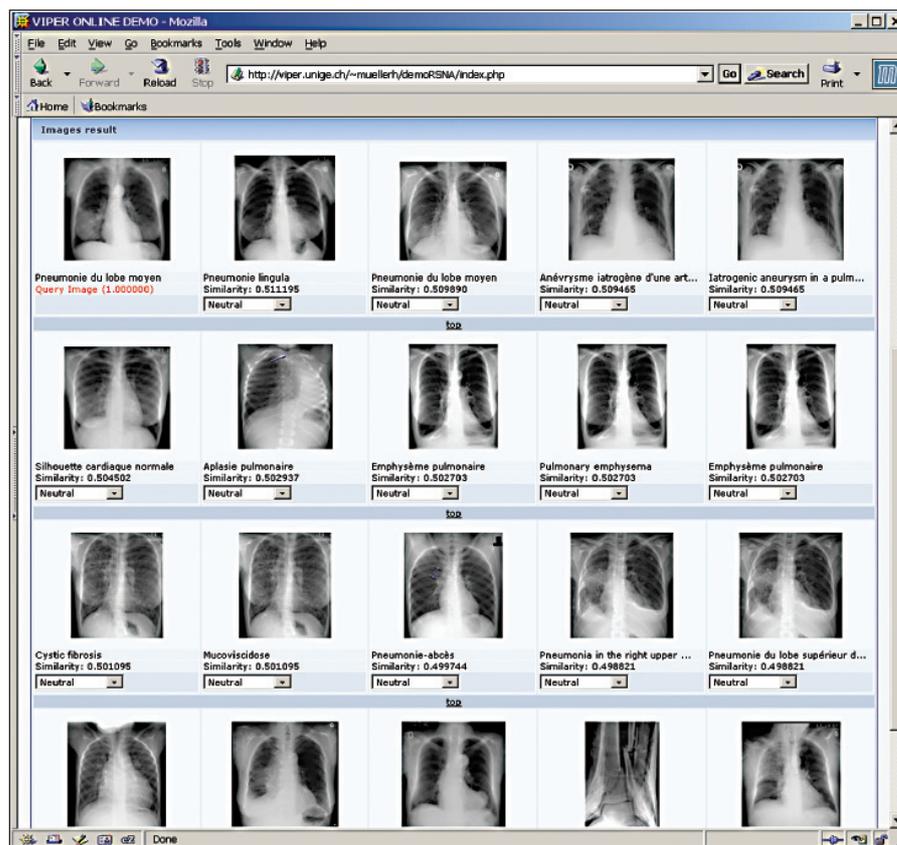


Abbildung 5: Das Interface eines typischen visuellen Suchsystems mit einem Anfragebild und den Resultaten nach visueller Ähnlichkeit sortiert.

4 Anforderungen und neue Forschungsrichtungen

Es gibt sicherlich eine Vielzahl von Methoden, die einen Einfluss auf die Entwicklung der inhaltsbasierten Suche haben können. Zum einen sind dies technische Systeme, wie z. B. der standardisierte Zugriff auf anonymisierte Daten aus der elektronischen Patientenakte, ohne für jedes Projekt eine neue Datenbank und Infrastruktur erstellen zu müssen. An der Uniklinik Genf wird daran bereits im europäischen Projekt AneurIST¹¹ mit insgesamt 29 weiteren europäischen Partnern geforscht, um in den nächsten Jahren eine entsprechende Struktur zu entwickeln und praktisch zu testen.

Ein weiteres Problem ist die fehlende Infrastruktur bezüglich Rechnerressourcen in Krankenhäusern: Anders als in Forschungsbereichen wie der Kernphysik oder der Wettervorhersage in denen der Einsatz von Hochleistungsrechnern üblich ist, laufen die Forschungsanwen-

dungen aufgrund fehlender Ressourcen auf einfachen Bürorechnern. Da aber gleichzeitig in den meisten Kliniken sehr viele Rechner (im Uniklinikum Genf gibt es im Moment etwa 6.000 PCs) vorhanden sind, und die meisten nur selten benutzt werden, liegen Konzepte wie grid computing nahe (computing on demand). Hierbei werden die Aufgaben parallelisiert und an die verteilten Rechner geschickt die gerade nicht benutzt werden. Leider sind die meisten grids im Moment noch relativ schwierig zu installieren und zu warten. Ein Projekt, das helfen soll dies zu verändern, ist KnowARC¹², das basierend auf der ARC (Advanced Resource Connector) Middleware des Nordugrid Konsortiums versuchen wird, eine einfach zu wartende Grid Infrastruktur zu erstellen. Wichtig ist, nur Teile der Ressourcen zu benutzen, damit die Rechner noch für andere Anwendungen verfügbar sind. Große Grid-Projekte wie EGEE¹³ (Enabling Grids for E-Science in Europe) benötigen meisten sämtliche Ressourcen eines Rechners.



Neben diesen technischen Systemen sind vor allem Anwendungen nötig, die den Benutzer mit in die Entwicklung einbeziehen und die Benutzer auf die Eigenheiten der Technik hinweisen, um Vor- und Nachteile besser verstehen zu können. Dies kann zum einen durch Prototypen geschehen, die den Nutzern zugänglich gemacht werden, aber auch durch Umfragen unter medizinischen Datenbenutzern, welche Techniken für sie nützlich sind, oder welche Anwendungsbereiche sie sich vorstellen können. Vor allem ist es im Moment wichtig, die Kommunikation zwischen Forschenden und potentiellen Anwendern zu verbessern, um einen Transfer der Techniken zu gewährleisten und zu verhindern, dass Prototypen auf zur Verfügung gestellten Testdaten erstellt werden, aber womöglich weder als Anwendung nützlich sind noch auf anderen Daten funktionieren.

Ein weiterer wichtiger Bereich ist das zur Verfügung stellen von Referenzbildern oder Bilddatenbanken, um Forschungsergebnisse auf derselben Grundlage vergleichen zu können [14, 20], und zu verhindern, dass immer aufs Neue Ressourcen für die Erstellung von Datenbanken benötigt werden. Am Besten wäre es, wenn Publikationen die Rohdaten mit enthalten müssten, damit andere Forschungsgruppen auf denselben Daten arbeiten könnten und somit eine Vergleichbarkeit von Ergebnissen gewährleistet ist.

5 Ausblick

Die Menge und die Komplexität der im medizinischen Bereich digital produzierten multimedialen Daten wird weiter zunehmen. Um das in den Daten gespeicherte Wissen im größtmöglichen Umfang zu nutzen, müssen Techniken entwickelt werden, die den Zugriff auf die Daten effektiver ermöglichen als eine einfache Suche nach Stichworten. Die inhaltsbasierte visuelle Suche hat das Potential eine solche wichtige Technologie für das Management multimedialer medizinischer Daten zu werden. Allerdings sind noch verschiedene wichtige Schritte dafür nötig. Der wichtigste ist sicherlich die effektive Verbindung von verschiedenen Medien zum Data Mining und Information Retrieval, anstelle der im Moment separat genutzten Techniken für jedes einzelne Medium.

Als zweites müssen die Anwender Erfahrungen mit den Techniken sammeln können, damit die visuelle Suche in wenigen Jahren, ähnlich wie zur Zeit die Stichwortsuche mit Google, von nahezu jedem zumindest in Ansätzen beherrscht wird.

Die inhaltsbasierte Suche hat sicherlich noch einen langen Weg vor sich, bevor sie als eine Standardmethode für das medizinische Informationsmanagement bezeichnet werden kann.

Referenzen

- 1 AAT Bui, RK Taira, JD Dioniso, DR Aberle, S El-Saden, H Kangarloo, Evidence-based radiology – requirements for electronic access, *Academic Radiology* 9:662-669, 2002.
- 2 AM Aisen, LS Broderick, H Winer-Muram, CE Brodley, AC Kak, C Pavlopoulou, J Dy, CR Shyu, A Marchiori, Automated storage and retrieval of thin-section CT images to assist diagnosis: System description and preliminary assessment, *Radiology*, 228:265-270, 2003.
- 3 ES Berner, JR Jackson, and J Algina, Relationships among performance scores of four diagnostic decision support systems. *Journal of the American Medical Informatics Association*, 3(3): 208-15, 1996.
- 4 M Bober, MPEG-7 Visual Shape Descriptors, *IEEE Transactions on Circuits and Systems for Video Technology*, 11(6):716-719, 2001.
- 5 C Carson, S Belongie, H Greenspan J Malik, Blobworld: Image Segmentation Using Expectation Maximization and its Application to Image Querying. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 1026-1038, 2002.
- 6 P Clough, H Müller, T Deselaers, M Grubinger, TM Lehmann, J Jensen, W Hersh, The CLEF 2005 Cross-Language Image Retrieval Track, *Springer Lecture Notes in Computer Science LNCS 4022:535-557*, 2005.
- 7 P Clough, M Sanderson, H Müller, The CLEF Cross Language Image Retrieval Track (ImageCLEF) 2004, *Springer Lecture Notes in Computer Science* 3115:243-251, 2004.
- 8 T Deselaers, D Keysers, H Ney, Features for Image Retrieval: A Quantitative Comparison, *Proceedings DAGM Tübingen* 2004.
- 9 T Deselaers, D Keysers, H Ney, Discriminative Training for Object Recognition using Image Patches, *Proceedings CVPR 2005*, 2:157-162, San Diego, 2005.
- 10 T Deselaers, D Keysers, H Ney, Improving a Discriminative Approach to Object Recognition using Image Patches, *DAGM 2005, LNCS 3663*, 2005.
- 11 PGB Enser, Pictorial information retrieval, (*Progress in Documentation*) *Journal of Documentation* 51(2):126-170, 1995.
- 12 R Fergus, P Perona, A Zissermann. Object Class Recognition by Unsupervised Scale-Invariant Learning, *Conference on Computer Vision and Pattern Recognition*, 264-271, Blacksburg, VA, USA, 2003.
- 13 RM Haralick, B Shammugam, I Dinstein, Texture Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6): 610-621, 1973.
- 14 A Horsch, M Prinz, S Schneider, O Sipilä, K Spinnler, JP Vallee, I Verdonck-de Leeuw, R Vogl, T Wittenberg, G Zahlmann, Establishing an International Reference Image Database for Research and Development in Medical Image Processing, *Methods of Information in Medicine* 43:409-415, 2004.
- 15 D Keysers, C Gollan, H Ney, Classification of Medical Images using Non-linear Distortion Models. *Bildverarbeitung für die Medizin*, 366–370, Berlin, Germany, 2004.
- 16 TM Lehmann, MO Güld, C Thies, B Fischer, K Spitzer, D Keysers, H Ney, M Kohnen, H Schubert, B Wein, Content-based image retrieval in medical applications, *Methods of Information in Medicine* volume 43(4): 354-361, 2004.
- 17 H Lowe, I Antipov, W Hersh, C Arnot Smith, Towards Knowledge-Based Retrieval of Medical Images. The role of semantic indexing, *Image Content Representation and Knowledge-Based Retrieval*, Annual Symposium of the American Society for Medical Informatics



- matics (AMIA), 882-886, 1998.
- 18 R Maree, P Geurts, J Piater, L Wehenkel. Random Subwindows for Robust Image classification, IEEE Conference on Computer Vision and Pattern Recognition, 1:34-40, San Diego, CA, USA, IEEE, 2005.
 - 19 H Müller, N Michoux, D Bandon, A Geissbuhler, A review of content-based image retrieval systems in medicine – clinical benefits and future directions, International Journal of Medical Informatics, 73:1-23, 2004.
 - 20 H Müller, A Rosset, JP Vallée, F Terrier, A Geissbuhler, A reference data set for the evaluation of medical image retrieval systems, Journal on Computerized Medical Imaging and Graphics, 28:295-305, 2004.
 - 21 C Palm, D Keysers, T Lehmann, K Spitzer, Gabor Filtering of Complex hue/Saturation Images for Color Texture Classification. Proc JCIS 2000, 45-49, Atlantic City, NJ, USA, 2000.
 - 22 H Qi and WE Snyder, Content-Based Image Retrieval in PACS, Journal of Digital Imaging 12(2), 1999.
 - 23 AR Rao, A Taxonomy for Texture Description and Identification, Springer 1990.
 - 24 A Rosset, O Ratib, A Geissbuhler, JP Vallée, Integration of a Multimedia Teaching and Reference Database in a PACS Environment, Radiographics, 22(6):1567-1577, 2002.
 - 25 P Schmid-Saugeon, J Guillod, JP Thiran, Towards a computer-aided diagnosis system for pigmented skin lesions, Computerized Medical Imaging and Graphics, 27:65-78, 2003.
 - 26 AWM Smeulders, M Worring, S Santini, A Gupta, R Jain. Content-based image retrieval: The end of the early years, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12):1349-1380, 2000.
 - 27 DM Squire, W Müller, H Müller, J Raki. Content-based query of image databases, inspirations from text retrieval: Inverted files, frequency-based weights and relevance feedback, Scandinavian Conference on Image Analysis, 143-149, Kangerlussuaq, Greenland, June 1999.
 - 28 HD Tagare, C Jaffe, J Duncan, Medical Image Databases: A Content-Based Retrieval Approach, Journal of the American Association for Medical Informatics (JAMIA), 4(3):184-198, 1997.
 - 29 H Tamura, S Mori, T Yamawaki, Textual Features Corresponding to Visual Perception, IEEE Transactions On Systems, Man, and Cybernetics 8(6): 460-472, 1978.

Fußnoten

- 1 <http://www.casimage.com/>
- 2 <http://www.rsna.org/>
- 3 <http://mirc.rsna.org/>
- 4 <http://www.healcentral.org/>
- 5 <http://www.pathopic.org/>
- 6 <http://elib.cs.berkeley.edu/blobworld/>
- 7 <http://www.mrml.net/>
- 8 <http://trec.nist.gov/>
- 9 <http://www.clef-campaign.org/>
- 10 <http://www.ir.shef.ac.uk/imageclef>
- 11 <http://www.aneurist.org/>
- 12 <http://www.knowarc.eu/>
- 13 <http://www.egee.org/>

Kontakt

Henning Müller (PhD)
*University & Hospitals of Geneva
Service of Medical Informatics
24, rue Micheli-du-Crest
CH-1211 Geneva 14
Switzerland
Tel.: + 41 (0) 22 3 72 - 61 75
Fax: + 41 (0) 22 3 72 - 86 80
henning.mueller@sim.hcuge.ch*