

AUTOMATISIERTE, INHALTSBASIERTE IMAGE-RETRIEVAL-SYSTEME

Zusammenfassung

Das Wachstum der Datenmengen und die Erleichterung der Zugriffsmöglichkeiten auf Multimedia-Datenbanken hat die Herangehensweise an die Informationsbeschaffung nachhaltig verändert. Klassische textbasierte Systeme haben im Kontext multimedialer Informationssuche ihre Grenzen erreicht. In diesem Text geben wir einen Überblick über die Arbeitsweise inhaltsbasierter, automatisierter Image-Retrieval-Systeme (Systeme zur Bildersuche bzw. zur Beschaffung visueller Informationen d. Übers.) und deren praktische Umsetzung.

- Wir beginnen mit der technischen Analyse von inhaltsbasierten Image-Retrieval-Systemen (CBIR-Systeme). Als praktisches Beispiel stellen wir unser System Viper vor, das an der Universität von Genf entwickelt wurde.¹ CBIR-Systeme sind spezifische Werkzeuge; sie zielen darauf hin, bestimmte Teile des menschlichen Wahrnehmungsapparates zu simulieren. Die Auswertung der Analyse muß diesem Umstand Rechnung tragen. Auf offene Fragen und Entwicklungsmöglichkeiten kommen wir am Ende zu sprechen.

Einführung

Das stetige Zunahme der Zugriffsmöglichkeiten auf multimediale Daten und Datenbanken über das World Wide Web (WWW) macht es nötig, Werkzeuge zu entwickeln, die eine effiziente Abfrage dieser Archive ermöglicht. Die Konzeption eines solchen Systems hängt von der Herangehensweise des Entwicklers ab und reicht von digitaler Bildbearbeitung und -analyse bis hin zu reiner Software-Entwicklung.

Es gibt bereits einige Bildersuchmaschinen im WWW. Entweder handelt es sich dabei um kommerzielle Anwendungen oder um Prototypen als Ergebnisse wissenschaftlicher Forschung, die in beigefügten Forschungsberichten erläutert werden. Obwohl die Systeme sich in Konzeption und Herangehensweise vielfältig unterscheiden, haben alle das gemeinsame Ziel, dem Anwender bei der Suche nach visueller Information in einer Datenbank zu unterstützen. Dabei muß man den Standpunkt des Benutzers einnehmen um von dort aus zu ermitteln, welches Maß an Unterstützung ein solches System unterschiedlichen Anwendern bereitstellen sollte. Das wiederum erlaubt es festzustellen, welches Verfahren für eine Suchanfrage nach Bildern eingesetzt werden sollte. Verfahren der

¹ Dieses Projekt wird vom Schweizerischen Nationalfonds zur Förderung der wissenschaftlichen Forschung gefördert (Projekt-Nr. 2000-052426.97).

Bildererkennung und -verarbeitung bilden den Kern inhaltsbasierter Image-Retrieval-Systeme (Content-Based Image Retrieval, kurz: CBIR), da sie festlegen, in welcher Repräsentationsform die Dokumente miteinander verglichen werden.

Während große Fortschritte in der Technik der Bildanalyse gemacht wurden bestehen bei der automatisierten, inhaltsbasierten Bildverarbeitung noch grundlegende Schwierigkeiten mit der Semantik eines visuellen Dokumentes. Ganzer Satz: Dieser Umstand übt einen starken Einfluß auf die interne Repräsentation des Bildes und folglich auch auf die Auswahl der Kriterien, mit denen nach bestimmten Merkmalen gesucht werden soll. Ebenso muß man die Architektur und das Zusammenspiel der verschiedenen Komponenten des gesamten CBIR-Systems unter dem Gesichtspunkt der Software-Entwicklung betrachten, um von dem Kontext, für den die Anwendungen entwickelt werden (hier das World Wide Web) zu profitieren.

Untersuchen wir also einige Aspekte des CBIR-Systems, von der Benutzerfreundlichkeit bis hin zur praktischen Auswertung. Unsere Untersuchung steht in engem Zusammenhang mit den Erfahrungen im Bereich textbasierter Suchmaschinen, also vergleichsweise ausgereifter Systeme. Wir betrachten das Problem der Entwicklung eines CBIR-Systems von verschiedenen Standpunkten.

Es ist notwendig, einen Ausgangspunkt zu definieren, der sich aus den Anforderungen ergibt, die der Nutzer, die Entwicklung und die Architektur an das System stellen. Um die Konstruktion eines IBIR-System darzustellen, verwenden wir als konkretes Beispiel das System Viper, das seit vielen Jahren in unserer Gruppe entwickelt wird. Abschließend behandeln wir die Probleme, die noch immer verhindern, daß CBIR-Anwendungen in lauffähigen Programmen wirkungsvoll eingesetzt werden können. Dieses Thema birgt noch viele offene Fragen.

Die inhaltsbasierte Abfrage von Bildern

Um den Zugriff auf die Daten einfach zu gestalten, ist es außerordentlich wichtig, die Datenabfrage so zu gestalten, daß sie der Intuition des Benutzers möglichst genau entspricht. Die gängigen Suchmaschinen benutzen eine textbasierte Abfrage. Der Benutzer wird aufgefordert die Suchbegriffe oder auch ganze Sätze und Satzteile einzugeben, die die gesuchte Information am besten umschreiben (z. B. AltaVista, 2000). Solch ein Abfragesystem setzt nicht nur voraus, daß in der abgefragten Datenbank korrespondierende Schlüsselwörter existieren, nach denen die Daten geordnet werden können, sondern auch, daß die gespeicherten Informationen mittels einer begrenzten Anzahl von Suchbegriffen ausreichend beschrieben werden können.

Das muß aber nicht der Fall sein, wenn es sich um visuelle Informationen handelt, wie es bei Bildern der Fall ist. Obwohl man den sichtbaren Gehalt visueller Information leicht durch die Benennung der Kompositionsteile beschreiben könnte, treten Schwierigkeiten auf, wenn man eindeutig und exakt die Untertöne

und den emotionalen Gehalt des Bildinhaltes beschreiben will. Und diese Schwierigkeiten werden größer, wenn der Benutzer nur eine vage Idee von dem hat, wonach er sucht.

Das Paradigma der Beispielabfrage (*query by example*, kurz QBE) stellt eine Lösung des oben genannten Problems dar. Im folgenden stellen wir die Charakteristika eines typischen QBE-basierten CBIR-Systems vor. Um die Voraussetzungen und Konsequenzen eines solchen Systems besser bewerten zu können, betrachten wir es unter verschiedenen Aspekten.

Suchparadigmen und Interaktivität

Gewöhnliche Abfragesysteme benutzen Suchbegriffe, um die gesuchte Information zu beschreiben. Das Entstehen multimedialer Datenbanken, insbesondere von Bildarchiven, hat gezeigt, daß diese Art der Abfrage im Umfeld multimedialer Informationssuche zu restriktiv ist. Die Ursache dafür liegt in der Tatsache, daß ein schriftlicher Kommentar nicht in jedem Fall flexibel genug ist, um Bildinhalte zu beschreiben oder um eine Abfrage zu formulieren, die alle für den Benutzer bedeutsamen Aspekte abdeckt. Darüber hinaus ist es möglich, daß — für jeden Nutzer subjektiv — eine bestimmte Kombination von Suchbegriffen, mit einem Dokument assoziiert wird, sei es in einem bestimmten Kontext oder von einem bestimmten Typ von Benutzern.

Das QBE-Paradigma legt nahe, die Formulierung der Abfrage an dem konkret untersuchten Datentypus, hier also der Bilder, auszurichten. Die Grundlagen des QBE-Paradigmas können folgendermaßen formuliert werden.

- Für den Benutzer ist es leichter, Beispiele für die gesuchte Information zu finden, als die gesuchte Information zu beschreiben.
- Die Abfrage ist gezielter und näher an der natürlichen Fragestellung des Benutzers.
- Die Ergebnisse können vom Benutzer durch direkten Vergleich mit der Ausgangsfrage ausgewertet werden.

Im Idealfall gibt der Benutzer ein Beispiel und das CBIR-System zeigt als Ergebnis alle Bilder der Datenbank auf, die der Abfrage entsprechen. Die meisten bestehenden Systeme bieten eine zufällige Auswahl von Beispielbildern aus ihrer Datenbank als eine Art von Ausgangspunkt für den Benutzer an. Da aber die meisten dieser Systeme nur ihre Fähigkeit demonstrieren wollen, ähnliche Bilder aufzufinden, hat der Benutzer nicht wirklich die freie Wahl eines Ausgangspunktes für eine Abfrage.

Eine mögliche Lösung dieses Problems wäre es, wenn der Benutzer seinen eigenen Ausgangspunkt skizzieren könnte (z.B. als Skizze oder Bildkomposition). Wenn man es allerdings dem Nutzer erlauben wollte, eigene Bilder als Ausgangspunkt zu nehmen, setzt das voraus, daß diese Bilder in irgendeiner Form zum Zweck des Vergleichs in die Datenbank aufgenommen werden. Da

dies lange Rechenzeiten in Anspruch nimmt, wird diese Möglichkeit nicht immer angeboten.

In jedem Fall ist es unwahrscheinlich, daß der Nutzer ein Beispielbild findet, das exakt alle Merkmale beinhaltet, die er sucht (wenn dem so wäre, hätte er schon gefunden, was er sucht). Es bietet sich also an, das QBE-Paradigma als Abfragesystem zu erweitern, bei dem der Nutzer mehrere Beispiele anführt und das CBIR-System die angeführten Informationen in einem Pseudo-Bild zusammenführt und dazu den besten Treffer in der Datenbank ausgibt. Man sollte festhalten, daß das QBE-Paradigma hier nicht das textbasierte Abfragesystem ersetzt, sondern erweitert (QBIC, 1998).

In Anbetracht der subjektiven Bewertung durch die Nutzer und der Unvollständigkeit der Datenbanken und also auch der Beispielkataloge (d.h. Diskontinuität der durchsuchten Menge), wäre es unrealistisch anzunehmen, daß das CBIR-System gleich im ersten Schritt der Abfrage das korrekte Ergebnis ausgibt. Man muß dem Nutzer also die Möglichkeit bieten, die Abfrage auf der Grundlage des ersten Ergebnisses zu präzisieren. Man könnte zum Beispiel die Bewertung der Suchergebnisse nach Wichtigkeit nutzen, um die Beispielabfrage zu generalisieren.

Ausgehend von den als Suchresultat ausgegebenen Bildern wird die Abfrage reformuliert, indem man relevante und irrelevante Ergebnisse markiert und so mit jedem Schritt Informationen betont oder ausklammert. Dadurch entsteht ein dynamischer Dialog zwischen dem Nutzer und dem System, der zielbewußt Lerneffekte simuliert (Müller et. al., 1999),(Müller et. al., 2000a),(Squire et al., 2001).

Bildrepräsentation

Das Hauptziel der Beispielabfrage ist es, die Abfrage intuitiver zu gestalten, um damit den Nutzer in eine angenehmere Ausgangslage zu bringen. Aus der Perspektive des Rechners stellen sich die Dinge ganz anders dar.

Das Paradigma der QBE basiert *qua definitionem* auf einer gründlichen Inhaltsanalyse des Abfragebildes. Dieses Paradigma muß also alle rechnerischen Möglichkeiten der Bilderkennung nutzen. An dieser Stelle wollen wir keine abschließende Lösung für diesen Teil des Systems anbieten, sondern die Alternativen betrachten und analysieren, die man normalerweise in der Literatur zur Bildverarbeitung findet. Die Analyse eines Bildinhaltes orientiert sich über die Definition von Bildmerkmalen, deren Repräsentation in einem Vektor gespeichert werden. Solch ein Eigenschafts-Vektor ist die Grundlage aller anschließenden Analysen. Bildmerkmale können von der Farbanalyse bis hin zur Beschreibung markanter Punkte, z. B. Eckpunkte reichen. Merkmale höherer Ordnung, z. B. die Form von Objekten oder eine lokale Krümmung von Linien können zur Beschreibung des Bildinhaltes genutzt werden.

Spezielle Anwendungsgebiete erfordern möglicherweise die Definition hochspezialisierter charakteristischer Repräsentationsformen für die Bildinhalte. Das gilt etwa für medizinische Visualisierungen wie die MRI-Abbildungen der Hirnregion, auf denen Laien kaum Unterschiede erkennen können. All diese Merkmale mögen im ganzen Bild oder aber in Teilbereichen extrahiert werden, wie sie durch Quadtree-Verfahren (sukzessive Viertelung in Baumstruktur) oder andere Segmentierungstechniken zustande kommen. Es gibt verschiedene Möglichkeiten, Bildcharakteristika für spätere Abfragen zu speichern, etwa durch Bündelung in einem Histogramm. Ein typisches Beispiel dafür ist das Farbhistogramm eines Bildes, das die Häufigkeit aller Farben (je nach ihrer Quantisierung) im Bild tabellarisch aufträgt. Diese Bilddarstellung hat den Vorteil, das gesamte Bild zusammenzufassen und gegenüber Rotationen invariant zu sein, was für manche Anwendungen hilfreich ist. Andererseits ist sie zu allgemein, um Bildinhalte präzise zu beschreiben. Diese Problem läßt sich lösen, indem man das Bild rekursiv unterteilt und dann Histogramme erstellt. In diesem Fall würde jeder Teil des Bildes nach eigenen Kriterien beschrieben. In jedem Fall ist es nötig, zwischen ortsunabhängigen (bzw. globalen) und speziellen (lokal begrenzten) Merkmalen der Bildrepräsentation abzuwägen.

Farbmerkmale

Farbe ist ohne Zweifel das erste Merkmal, an das man denkt, wenn man ein Bild charakterisieren will. Mit der Definition des Farbraumes, der jede mögliche Farbe umfaßt, sind die Fragen des Lichteinfalles und der Perspektive ausgeklammert. Definitionen der grundlegenden Farbräume finden sich in jedem Standardwerk zur Bildverarbeitung (Jain 1989). Beleuchtungsunabhängige Farbräume haben Gevers und Smeulders (1996) gründlich untersucht.

Texturmerkmale

Im Kontext multimedialer Prozesse versteht man unter Textur das Maß der Gleichförmigkeit innerhalb eines Bildausschnitts. Die Textur eines Bildteils zeigt sich als Struktur einer Wiederholung oder einer bestimmten Anordnung. (z. B. Leinwand, Holz, Mauerwerk). Mittlerweile gilt die Textur als entscheidende Charakteristik jeder inhaltlichen Bildanalyse. Die Untersuchung der statistischen Verteilung solcher strukturellen Merkmale sollte es ermöglichen, jede Textur in ihrer Art (also grob, fein, regelmäßig etc.), eindeutig zu beschreiben.

Spezielle Merkmale

Für spezifische Anforderungen kann es hilfreich sein, sehr spezialisierte Merkmale zu entwickeln, die auf vorhandenen Kenntnissen basieren. Beispiele spezialisierter Merkmale sind Umrißdeskriptoren oder Filter zur Lokalisierung von Bildteilen.

Je nach betrachtetem Datentyp kann der Umriss unter bestimmten Umständen solch ein Kriterium sein. Wenn alle in einem Bild enthaltenen Objekte zuerst isoliert werden sollen, muß das Bild in irgendeiner Form unterteilt werden. Die Objekte werden einzeln betrachtet und ihre Merkmale werden von Eigentümlichkeiten des Objekts oder dessen Verhältnis zur Bildumgebung hergeleitet. Manchmal sind auch Merkmale von Nutzen, die aus einem kontrollierten Lernprozeß gewonnen werden, z.B. Parametermuster, die durch den Vergleich verschiedener Beispielbilder gewonnen werden bei der Erkennung von Gesichtern (Rowley et al. 1998) Diese technologische Entwicklung stieß besonders bei Entwicklern forensischer und sicherheitstechnischer Programme auf großes Interesse. Normalerweise werden hier Porträtfotografien in neuronale Netze eingespeist. Das neuronale Netz wird dann zum Detektor, der jedesmal „anschlägt“, wenn er in einem Bildbereich ein spezifisches Objekt (in diesem Fall ein menschliches Gesicht) vermutet.

Systemkomponenten und Organisation

Die Entwicklung eines CBIR-Systems beinhaltet nicht nur die Suche nach einem möglichst guten Abfragesystem und einer geeigneten Bildrepräsentation. Da es große Datenmengen verarbeiten muß, sollte es wie ein komplexes System betrachtet werden. Im Folgenden wollen wir die wichtigsten Komponenten skizzieren, die diesen Aspekt beim Entwurf eines CBIR-Systems betreffen.

Aus dem bisher gesagten wurde deutlich, daß die Struktur eines CBIR-Systems sich in drei Teile aufteilt: die Benutzerschnittstelle, die Suchmaschine und das Verfahren der Indizierung. Jeder dieser Teile sollte als eigenständige Komponente analysiert werden, so daß er bei Bedarf ausgetauscht oder aber in anderen Systemen angewendet werden kann.

Die Benutzerschnittstelle

Bei der Benutzerschnittstelle handelt es sich um den sichtbaren Teil des CBIR-Systems. Um diesen Teil der Software zu entwickeln, sind Kenntnisse auf dem Gebiet der Mensch-Computer-Interaktion hilfreich. Zweifellos ist das zuwenig beachtet worden, weil die Schnittstelle in das System integriert ist und deshalb von denselben Entwicklern realisiert wird, wie der Rest des Systems. Wir sind davon überzeugt, daß die Benutzerfreundlichkeit der Schnittstelle ein wichtiger Faktor für die Qualität des ganzen Systems ist. Auch wenn das System großartige Möglichkeiten anbietet, kann der Nutzer sie oft nicht anwenden, weil sie in der Oberfläche nicht dargestellt werden und deshalb ohne Expertenwissen verborgen bleiben.

Die Bildersuchmaschine

Das Herzstück des CBIR-Systems ist die Bildersuchmaschine. Dieser Teil des Systems verarbeitet die Abfrage des Nutzers und gibt ein Suchergebnis aus. Die Hauptaufgabe dieses Systemteils ist es, gespeicherte Information zu suchen und dabei Ähnlichkeitsmaße einzusetzen, die möglichst nahe Treffer finden. Um die Aufmerksamkeit des Nutzers nicht zu strapazieren, sollte die Rechenleistung während der Abfrage minimiert werden, um möglichst rasch eine Antwort zu liefern. Zu diesem Zweck ist es entscheidend, möglichst viele Rechenschritte in die Vorverarbeitung, also die Indizierung der Daten, zu verlagern.

Der Bildindizierung

In diesem Teil des Systems werden die Daten vorverarbeitet, damit sie möglichst schnell gesucht und gefunden werden können. Bei der Indizierung der Bilder werden all die Aufgaben der Bilderkennung erledigt, die am meisten Rechenzeit kosten.

Diese Systemkomponente sollte erweiterbar zu sein, so daß bei der Indizierung weiterer Bildrepräsentationen ihre Struktur nur vervollständigt und nicht gänzlich neu berechnet werden muß.

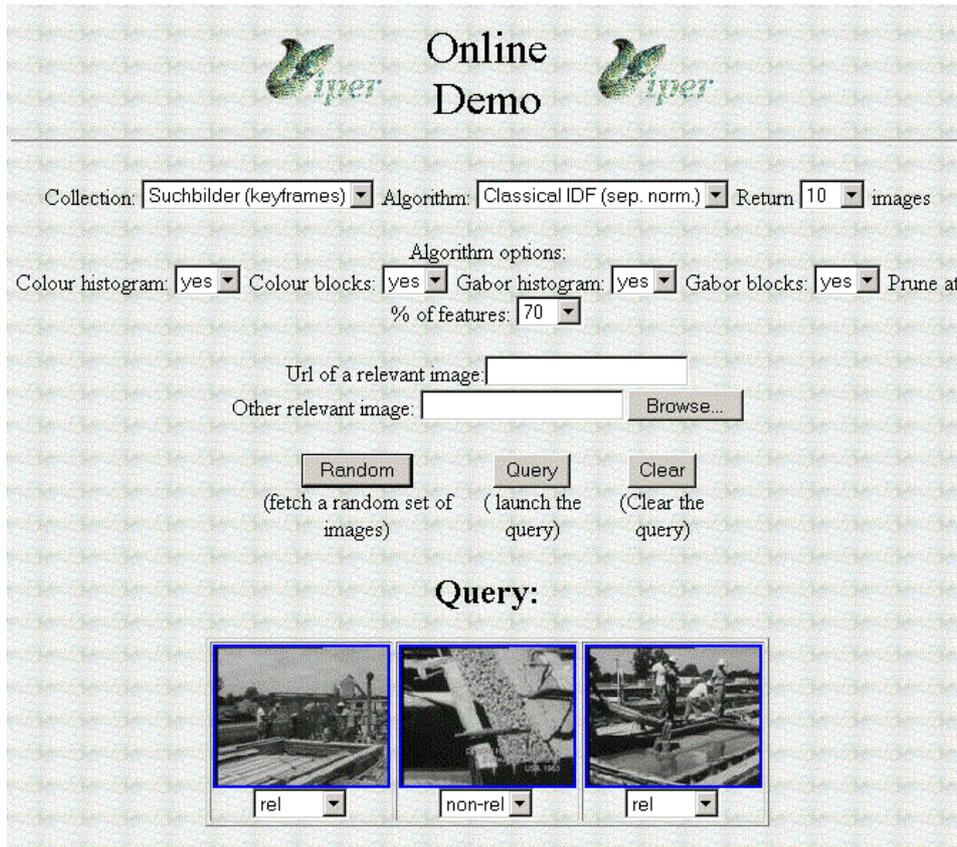
VIPER: Ein lauffähiges CBIR-System

Das CBIR-System Viper wurde von der Arbeitsgruppe „Computer Vision“ an der Universität von Genf entwickelt. Es basiert auf dem Suchparadigma der Beispielabfrage (*query by example*, QBE) und hat folgende Haupteigenschaften:

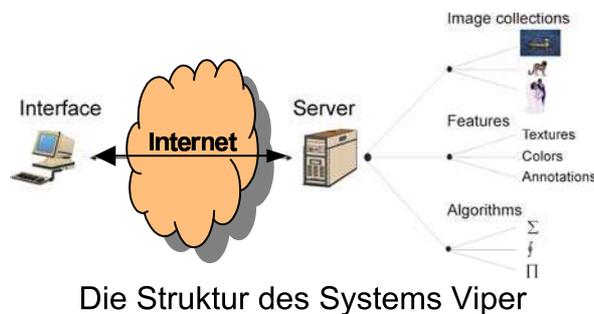
Das Suchsystem wurde weitgehend von Text-Suchsystemen beeinflusst. Insofern benutzt es einen sehr großen Katalog einfacher Bildmerkmale (und nicht etwa einen kleinen Katalog hochspezialisierter Bildkriterien).

Viper fragt über mehrere Schritte vom Nutzer eine Bewertung der Ergebnisse ab (*relevance-feedback*) und erlaubt damit eine zielgerichtete Suche.

Viper hat zwei Ziele. In erster Linie wird es als Entwicklungsplattform für ein ausbaufähiges, generisches CBIR-System genutzt. In diese Hinsicht ist es dafür ausgelegt, ohne Schwierigkeiten neue Kriterien und Suchalgorithmen integrieren zu können. Deswegen sollte es sehr einfach in hochspezialisierten Arbeitsgebieten, wie der medizinischen Bildsuche, einzusetzen sein.



Die Schnittstelle des Viper CBIR-Systems (<http://viper.unige.ch>)



Die Struktur des Systems Viper

Bildkriterien

Nach dem Vorbild textbasierter Suchsysteme benutzt das Viper-System eine große Anzahl einfacher Bildkriterien. Die aktuelle Version wertet lokal wie auch global Farb- und Frequenzmerkmale aus, die in unterschiedlicher Genauigkeit extrahiert werden, und nutzt deren Häufigkeitsverteilung sowohl am einzelnen Bild als auch in der kompletten Datenbank. So sollen einfache Bildeigenschaften, die (mehr oder weniger) der menschlichen Wahrnehmungsweise entsprechen, vom System aus verfügbar sein.

Das System bietet über 80.000 Merkmalsbeschreibungen. Jedes Bild besitzt $O(10^3)$ Merkmale, wobei die in den Bildern erfaßten Merkmale in einer invertierten Datei gespeichert werden. In Verbindung mit der Gewichtung der einzelnen Merkmale (s.u.) sorgt diese Datenstruktur dafür, daß schriftliche Anmerkungen ohne Probleme integriert werden können, weil sie genau so wie visuelle behandelt werden.

Erweiterte Interaktivität durch Rückkopplung von Bewertungen

An Stelle des Versuchs, aus den Rohdaten der Bilder semantische Elemente zu gewinnen, verfolgt Viper die Idee, aus der flexiblen Nutzer-Interaktion Informationen über den semantischen Inhalt der Suchbilder abzuleiten. Wenn die Merkmale mit den positiv oder negativ bewerteten Beispielbildern und deren Eigenschaften verknüpft werden, ergeben sich dynamische Beziehungen, die den Wünschen und Bedürfnissen eines Nutzers entsprechen. Zusammenfassend läßt sich sagen, daß Viper das Grundprinzip verfolgt, den Nutzer das tun zu lassen, was das System von alleine nicht leisten kann (z.B. die Formulierung eines semantischen Konzepts).

Über den Nutzen von CVIR-Systemen

Wie gezeigt ist es möglich, ein System zur inhaltsbasierten Bildersuche zu implementieren. Dieses System kann in Hinsicht auf technische Aspekte wie Geschwindigkeit, Kapazität und Anbindung des Speichers und der Architektur (z. B. wenn es im Internet eingesetzt wird) noch optimiert werden. Aber man muß sich die Frage stellen, ob solch ein System für praktische Anwendungen von Nutzen ist und wenn ja, für welche?

Voraussetzungen

Von vorne herein ist es offensichtlich, daß die wissenschaftlichen Anforderungen nicht unbedingt den Bedürfnissen des Nutzers entsprechen. Anders gesagt: wissenschaftliche Zielsetzungen erweisen sich im wirklichen Leben oft als sinnlos.

In gewisser Hinsicht gilt dies auch für das Paradigma der Beispielabfrage (QBE). Ist die Annahme vernünftig, der Nutzer könne seine Suche anhand von Bildbeispielen formulieren? Vielleicht würde er viel lieber ein eigenes Beispiel einsetzen. In diesem Fall müßte das Problem, ein reales Bild mit einer Skizze zu vergleichen, gelöst werden. Ferner werden im Kontext der CBIR-Bilder aber meistens als Träger von Information betrachtet, der künstlerische Aspekt wird dabei ausgeklammert. Wäre es möglich, diesen Aspekt in das System zu integrieren? Wäre es denkbar, das Charakteristische an einem „Van Gogh-Stil“ oder dem „Impressionismus“ zu definieren? Falls ein System ein künstlerisches Konzept nachvollziehen kann, sollte es dazu in der Lage sein, weitere Exemplare nach diesem Muster zu reproduzieren, oder nach dem Muster bekannter Texturen

weitere errechnen zu können. Das Ergebnis wäre dann tatsächlich computergenerierte Kunst.

Kontext

Ein anderer wichtiger Punkt betrifft die Machbarkeit von nützlichen CBIR-Systemen. Meistens wird davon ausgegangen, daß die Suche in einer begrenzten Menge von Bildern durchgeführt wird. In dieser Auswahl muß das gesuchte Bild nicht enthalten sein, vielleicht sogar nicht einmal die richtigen Beispiele für die Suche. Um den Nutzen des Systems zu garantieren, müßte man voraussetzen, daß die betrachtete Bildermenge einen sehr großen Bereich lückenlos abdeckt. Nur so läßt sich über die ganze Menge eine Art von „Kontinuität“ bildlicher Ähnlichkeit erreichen. Das Feedback relevanter Suchergebnisse über mehrere Abfragen ist auf eine lückenlos und kontinuierlich bestückte Bildermenge nicht angewiesen, sondern kommt auch mit lückenhaften Sammlungen zurecht. Hat man also im gesamten Bildraum einmal einen Bereich von Interesse entdeckt, so erschwert es die Technik des Bewertungs-Feedback sowohl, diesen Bereich näher zu erkunden, wie auch, ihn wieder zu verlassen.

Der Ausweg aus dieser Misere besteht in der Einführung spezieller Verfahren in Bereichen mit einem implizit feststehenden Kontext (wie z. B. medizinischen Aufnahmen). Dort benötigt man die Fachkenntnisse von Experten, um die Bilder differenziert zu betrachten (z. B. pathologische Befunde bei Lungenaufnahmen). Das würde zwar die Entwicklung spezialisierter Systeme fördern, aber es widerspräche der Idee, die Suchfunktion zu automatisieren.

Anwendungen

Auch wenn konkrete Anwendungen für CBIR-Systeme noch nicht genau feststehen, wird doch deutlich, daß das Management wachsender Mengen digitaler Information sie erforderlich macht. Bislang legen die begrenzten Erfahrungen mit CBIR-Systemen die Vermutung nahe, daß sie eher zusätzliche Funktionen innerhalb von Prozessen übernehmen, die von Menschen ausgeführt werden, als diese zu ersetzen.

Desktop-Anwendungen erfordern häufig ein Suchwerkzeug, das die Bearbeitung großer Datenmengen, wie in unserem Fall der Bilder, erleichtert. Solche Anwendungen scheinen das zukünftige Einsatzgebiet von CBIR-Systemen zu bilden.

Zusammenfassung

In diesem Aufsatz haben wir die wichtigsten Aspekte erörtert, die bei der Entwicklung eines inhaltsbasierten Image-Retrieval-Systems beachtet werden

müssen. Von der Herangehensweise an das Problem hängen unterschiedliche Ansätze und Lösungen ab.

Das erklärt, warum bereits existierende Systeme große Unterschiede aufweisen, was die Schnittstellen, die bildlichen Repräsentationen und die Suchfunktionen betrifft. Wir haben unser System Viper als Beispiel eines CBIR-Systems beschrieben. Die intensive Rückgriff auf Nutzer-Feedback macht unser System sehr flexibel und für unterschiedliche Kontexte anpassungsfähig. Die Weiterentwicklung des Systems wird an der Universität von Genf vorangetrieben. Der Leser wird diesbezüglich auf die Demo-Version des Systems verwiesen, die online geschaltet ist.

Prototypen von CBIR-Systemen existieren als kommerzielle Produkte oder wissenschaftliche Demoversionen. Zuletzt haben wir den Nutzen eines solchen Instruments in praktischer Hinsicht erörtert. CBIR-Systeme sind erst seit kurzem Gegenstand der Forschung und so fehlt es noch an Feedback zwischen wissenschaftlichen Entwicklern und tatsächlichen Nutzern. Das liegt teilweise an der Tatsache, daß die Anforderungen an die inhaltsbasierte Bildsuche (CBIR) weder wirklich begriffen noch formalisiert sind. Der rasante Fortschritt auf diesem Gebiet weist darauf hin, daß es schon bald effektive und nützliche Verfahren und Systeme im Bereich der inhaltsbasierten Image- (und Multimedia-)–Suchmaschinen geben wird.

Quellen

(AltaVista, 2000) Altavista search engine. <http://www.altavista.com>, 2000

(Gevers and Smeulders, 1996) T. Gevers und A. W. M. Smeulders, A comparative study of several color models for color image invariants retrieval, in: Proceedings of the First International Workshop ID-MMS'96, 17–26, Amsterdam, August 1996

(Müller et al., 2000) H. Müller, W. Müller, S. Marchand-Maillet, D. McG. Squire und T. Pun, Learning features weights from user behaviour in Content-Based Image Retrieval, in: ACM SIGKDD, International Conference on Knowledge Discovery and Data Mining, Boston, August 2000

(Müller et al., 2000a) H. Müller, W. Müller, D. McG. Squire, S. Marchand-Maillet und T. Pun, Strategies for positive and negative relevance feedback in image retrieval, in: International Conference on Pattern Recognition (ICPR'2000), Barcelona, September 2000

(Müller et al., 1999) W. Müller, D. McG. Squire, H. Müller und T. Pun, Hunting moving targets: an extension to Bayesian methods in multimedia databases, in: Panchanathan et al. (SPIE Symposium on Voice, Video and Data Communications,, 1998)

(QBIC, 1998) QBICTM – IBM's Query By Image Content.
<http://www.qbic.almaden.ibm.com/qbic/>, 1998

(Rowley et al., 1998) H. A. Rowley, S. Baluja and T. Kanade, Neural-network-based face detection, in: IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-20(1), 1998, 23–38.

(Squire et al, 2001) D. McG Squire, H. Müller, W. Müller, S. Marchand-Maillet und T. Pun, Design and Evaluation of a Content-based Image Retrieval System, in: Syed M. Rahman (Hg.), Design and Management of Multimedia Information Systems: Opportunities and Challenges, Idea Group Publishing, 2001

(Witten et al., 1994) I. H. Witten, Al. Moffat und T. C. Bell, Managing gigabytes: compressing and indexing documents and images, Van Nostrand Reinhold, 115 Fifth Avenue, New York, NY 10003, USA, 1994.