

Proseminar Multimedia Datenbanken und Retrieval QBIC

Venera Fefler
Institut für Web Science and Technologies
Universität Koblenz-Landau, Campus Koblenz

Zusammenfassung: Inhaltsbasiertes Multimedia Retrieval ist ein sehr aktuelles Thema, das mit den Suchtechniken von Digitalobjekten nach deren Inhaltsmerkmalen beschäftigt. In dieser Ausarbeitung wird nur auf die Bilddateien eingegangen. Die Arbeit soll den Überblick verschaffen, welche Suchmöglichkeiten in einer Bilddatenbank es gibt und wie die inhaltsbasierte Anfragen helfen die Ergebnisse aus den großen Datenmengen einzugrenzen.

Keywords: QBIC, inhaltsbasierte Bildersuche, Multimedia Retrieval, Bilderdatenbank, IBM DB2

1. Motivation

In den letzten Jahren wachsen die Digitalbildsammlungen immer schneller und werden dabei immer größer und unübersichtlicher. Die Bilder werden überwiegend textbasierend gesucht. Die Bilddateien haben aber viel mehr Informationen zu bieten, als nur Dateiname oder die Schlüsselwörter. Die Suche, die sich nach Inhalt basiert, kann zu besseren Ergebnissen führen. Viele haben sich schon mit dem Thema beschäftigt.

Schon in den 90-er Jahren haben einige Forscher angefangen dieses Thema zu analysieren. Dennoch ist die inhaltsbasierte Bildersuche im Vergleich zum Fortschritt in den Bereichen Speicherung und Übertragung von Digitaldaten noch im Frühstadium.

Es gibt zwar einige kommerzielle Lösungen, die solche Funktionalitäten anbieten, sind aber immer noch nicht sehr verbreitet. Unter anderen sind folgende Produkte zu erwähnen: VIR Image Engine (Virage Incorporated), Excalibur Visual RetrievalWare (Excalibur Technologies Corporation), Photobook (Vision and Modeling Group, MIT Media Lab). Auch Forschungsteams an den Universitäten beschäftigen sich mit dem Thema.

Allerdings bleibt das Thema immer noch sehr aktuell.

2. Einführung

Ein Bild sagt mehr als tausend Worte". Dieses Sprichwort trifft genau das Ziel. Die Bilder spielen sehr große Rolle in unserem Leben. Es ist manchmal sehr schwer mit den Worten so genau zu beschreiben, was man mit den Augen sieht. Die visuelle Information wird schneller und besser verarbeitet als langer Text. Dieses Phänomen ist als Bildüberlegenheitseffekt[Klritzenberger] bekannt. Natürlich wird dies in der Informationsbranche sehr geschätzt und sich zunutze gemacht. Die wachsende Anzahl von digitalen Medien muss gut verwaltet werden, damit in den großen Ansammlungen noch eine effektive Suche möglich ist.

2.1 Wie werden die Bilder gesucht?

Die Bilder werden genauso wie auch die Daten im alphanumerischen Format indiziert in der Datenbank abgelegt. Leider sind die Möglichkeiten der Suchformulierung einer Bildersuche in einem 'normalen' Datenbanksystemen nur begrenzt. Es kann nur nach textuellen Eigenschaften der Datei selektiert werden, z.B.

Dateiname, Autor, Schlüsselwörter. Je nach Anwendungsfall, müssen die Bilder aber unterschiedlich gesucht werden.

Es gibt zwei Suchvarianten:

- explizit über die in der DB vorhandenen Daten
- impliziert über Multimedia-Retrieval-System

Der erste Fall wird verwendet bei Datenbankanwendungen, wenn eine Datei z.B. nach Name gesucht wird;

Der zweite Fall kommt in Einsatz, wenn nur 'besondere' Abbildungen in Betracht kommen können oder um die Suche einzugrenzen. Es handelt sich um die inhaltsbasierte Suche.

3. Was ist inhaltsbasierte Suche?

Inhaltsbasiert (content based) bedeutet eine Analyse des aktuellen Inhaltes eines Bildes, also der Farben (Color), Umrisse (Shape), Oberflächen (Texturen) oder anderer Informationen, die über eine automatische Bildverarbeitung ermittelt werden können.[wikipedia]

Ein großes Problem der inhaltsbasierten Suche ist die Formulierung der Suchkriterien. Die Anfragesprache soll die Kombination verschiedener Suchparadigmen vereinheitlichen. Zudem muss die Anfrageformulierung auch benutzerfreundlich sein, damit auch branchenfremde Benutzer solche Systeme nutzen können. Für die Anfrage werden folgende Aspekte berücksichtigt:

- Retrieval-Techniken ermöglichen das Abschätzen der Ähnlichkeit der Objeten
- Datenbankanfragetechniken ermöglichen auch auf statisch in der Datenbank abgespeicherten Informationen zuzugreifen
- XML-Anfragetechniken, da viele Multimediadaten lassen sich als xml-Daten darstellen
- Nutzerschnittstelle: Sowohl bei der Anfrage als auch bei der Ergebnisausgabe müssen spezielle Interaktionsschnittstellen verfügbar sein.

Um so ein Suchsystem zu erstellen, werden bestimmte Verfahren benötigt, die in der Lage sind, die relevanten Informationen aus den Digitaldateien zu extrahieren und auch zu vergleichen.

3.1 Die Suche nach Farben

Farbe ist eine der am öfteren verwendeten Eigenschaften des Bildes. Es ist unabhängig von der Bildgröße und Orientierung. Dabei ist es nicht nötig anzugeben, was genau da abgebildet ist. Es ist absolut unwichtig, ob es um ein Porträt, eine Landschaft oder eine technische Abbildung handelt. Entscheidend ist nur die Dominanz der eingegebenen Farben.

Als Farbeigenschaften benutzt QBIC :

- RGB-Farbraum
- Lab-Farbraum
- Yic-Farbmodell
- Munsell-Farbsystem
- 256-Farbhistogramm

Es werden die Durchschnittsfarbe und das Histogramm errechnet.

Um die Durchschnittsfarbe zu ermitteln, werden die Farbwerte aller Pixel aufsummiert und durch die Anzahl der Pixel geteilt. Mehrfarbige Flächen können dadurch die Durchschnittsfarbe haben, die gar nicht im Bild vorkommt.

Farbhistogramme stellen eine Häufigkeitsverteilung der in einem Bild vorkommenden Farben auf. Im Histogramm fehlen allerdings räumliche Informationen, so dass bei der Farbsuche dies gar nicht berücksichtigt werden kann.

D.h., dass es zwei absolut unterschiedliche Bilder ähnliche Histogramme liefern würden. Beispielweise, ein Schiff im Meer und ein Flugzeug im Himmel.

3.2 Die Suche nach Textur

Textur, oder Oberflächenstruktur, wird aus 3 Werten ermittelt

- Kontrast auf der Oberfläche
- Granularität, also die Größe sich wiederholenden Objekten
- dominierende Richtung der Abbildung

Die Textur wird als Grauwertmatrix dargestellt. Die gibt an, mit welcher relativen Häufigkeit zwei Grautöne in einer bestimmten Anordnung in einem Bild vorkommen.

3.3 Die Suche nach vorhandenen Konturen/Formen

Die Erkennung von Formen auf Bildern ist ebenso von großer Bedeutung bei QBIR. Die Suche nach Konturen setzt eine lokale Konturenbeschreibung voraus. D.h. welche geometrische Formen es im Bild vorkommen können, bzw. mit welchen vom User eingegebenen Konturen es verglichen werden kann.

Für diese Suche wird "Teplate-Matching-Methode" verwendet.

Unter Template Matching versteht man den Vergleich ("*Matching*") eines in Rasterform vorliegenden Prototypen eines Musters ("*Template*") mit einem zu untersuchenden Rasterbild. Für jede Position im Rasterbild wird die Korrelation (lat.: wechselseitige Beziehung) zwischen dem Prototyp und dem korrespondierenden Bildbereich ermittelt. Die Zuweisung oder Zurückweisung des Musters hängt von der Korrelation ab, d.h. von der Qualität des Vergleichs. Oder anders formuliert - nach der Berechnung der Korrelation wird entschieden, ob der Bildbereich dem Muster zugewiesen oder zurückgewiesen wird. Dies hängt von einem definierten Schwellenwert ab, der die Zuweisungsklasse von der Zurückweisungsklasse trennt. Template Matching eignet sich sehr gut für Einzelzeichen, Signaturen, etc. Akzeptable Rechenzeiten werden durch die Implementierung von Bildpyramiden erreicht. [TemplMatch]

Die Suche nach Konturen ist viel komplexer als die Suche nach Farben. Erstens das Erkennen einer Form auf dem betrachteten Bild. Zweitens muss diese Form in irgendeiner Art vorgehalten werden. Und drittens muss für Formen ein Maßstab entwickelt werden, um Bilder vergleichbar zu machen.

Im nächsten Abschnitt wird das IBM-Projekt QBIC als eine Lösungsvariante vorgestellt.

4. QBIC

Query By Image Content[ibm] ist eine kommerzielle Bildersuchtechnologie und Bestandteil des IBM DB2-Datenbanksystems. Die gehört zu den ersten Systemen, die nicht nur die Schlüsselwörter von Bilddateien, sondern auch deren visuellen Inhalt durchsucht haben. Mit dieser Technologie kann der Benutzer die Suchanfragen nach beinhaltenden Farben, Beispielbilder oder von ihm skizzierten Konturen stellen.

Damit solche Anfragen überhaupt gestellt werden konnten, müssen natürlich die abzufragende Informationen in der Datenbank vorhanden sein.

Schon beim Einfügen des digitalen Objektes in die Datenbank werden die einzelne Eigenschaften extrahiert und festgehalten. Diese Eigenschaften sind nichts anderes als die algorithmisch ermittelte mathematische Werte, mit denen die Ähnlichkeitsberechnung durchgeführt wird.

Diese Werte werden im sogenannten QBIC-Katalog abgelegt.

Dieser Katalog ist eine Art Datenbank, welche nicht nur Bilder, sondern auch alle extrahierten Daten über die visuellen Eigenschaften eines Bildes enthält. Jeder

6 Proseminar Multimediatdatenbanken und Retrieval: QBIC

Katalog besteht aus mehreren Dateien und kann mit weiterer alphanumerischer Information verknüpft werden.

Der DB2 Image-Extender unterstützt vier folgenden Merkmalsklassen[Pittner]:

Farbe:

- *QbColorFeatureClass(AverageColor)* berechnet die Durchschnittsfarbe eines Bildes
- *QbColorHistogramFeatureClass (Histogram)* erstellt von einem Bild ein Farbhistogramm

Textur

- *QbTextureFeatureClass (Texture)* misst die Granularität (Größe, der sich wiederholenden Elemente), den Kontrast (Helligkeitsvariationen) und die Gerichtheit (dominierende Richtung) der Textur des Bildes

Lage

- *QbDrawFeatureClass (Draw)* extrahiert den Durchschnittsfarbwert einer definierten Fläche

Für die Suche wird die Anfrage an den QBIC-Katalog gestellt. SQL-Anfrage benutzt folgende vier benutzerdefinierte Funktionen :

- QBScoreFromStr
- QBScoreFromName
- QBScoreTbFromStr
- QBScoreTbFromName

Die Funktionen liefern zurück eine reelle Zahl zwischen 0 und unendlich. Dieser Wert heißt Score und beschreibt den Abstand zum übergebenen Referenzbild. Dabei steht 0 für maximale Ähnlichkeit.

Eine Abfrage ordnet jedem Bild eine Bewertungszahl zu, die darüber Auskunft gibt, wie weit das gesuchte Objekt mit dem aktuellen Bild übereinstimmt. Ein hoher Wert kennzeichnet eine hohe Übereinstimmung, ein niedriger Wert entsprechend eine niedrige Übereinstimmung. Bei einer Abfrage über mehrere Bilder entsteht so ein Ranking, bei dem die interessantesten Bilder die höchsten Bewertungszahlen haben und damit am weitesten oben stehen.

Eine SQL-Anfrage sieht folgendermaßen aus:

```
SELECT
CONTENTS(image),
QBScoreFROMStr('QbColorFeatureClass=<255,0,0>',image) AS SCORE
FROM signs ORDER BY SCORE
```

QBIC Technologie ist Bestandteil verschiedener IBM-Produkte.

Zur Zeit wird die QBIC für die Online-Galerie des Hermitage Museums[herm] genutzt. Für Suchanfragen stehen zwei Möglichkeiten zur Verfügung: die Suche auf der Basis eines Vorgabebildes oder die Suche nach selbst gewählten Farben. In dieser Anwendung ist es nicht möglich, die Skizze selbst zu zeichnen, wobei das IBM-System diese Funktionalität anbietet. Bei der Layout-Suche auf der Seite wählt man die geometrische Figuren und fügt die dem Anfragebild zu.

An den folgen Ausschnitten von Hermitage Webseite wird die Funktionalität anhand der Beispielen noch mal dargestellt.

Die Suche nach Farbe. Man wählt die gewünschte (gesuchte) Farbe(n) von der Farbpalette und fügt die rechts hinzu. Beim Starten der Anfrage werden die beinhaltende Farbwerte ermittelt und die Ähnlichkeitsberechnung durchgeführt. Die Ergebnisse werden in absteigender Reihenfolge der Ähnlichkeit ausgegeben.

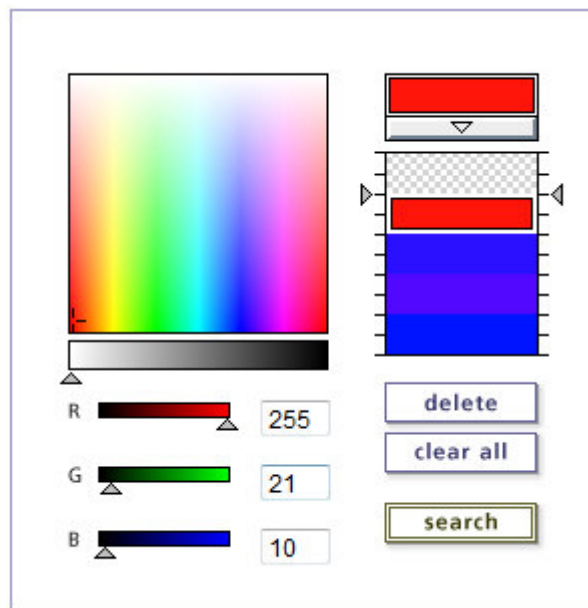


Abb1. Die Suche nach Farbe

8 Proseminar Multimediatdatenbanken und Retrieval: QBIC

So bekommt man dann die Ergebnisliste mit den Bildern, die Ähnlichkeit mit diesem Histogramm haben. Der folgende Ausschnitt zeigt die ersten vier relevanten Bilder.

SEARCH RESULTS

	1) <u>Boats at Saintes-Maries</u> Gogh, Vincent van 1888		2) <u>Portrait of Suzanne Dufy, the Artist's Sister</u> Dufy, Raoul 1904
	3) <u>Dance</u> Matisse, Henri 1909 - 1910		4) <u>Music</u> Matisse, Henri 1910

Abb2. Die ersten vier relevanten Bilder

Die Suche nach Layout. Hier wird die Suche nach vorhandenen im Bild geometrischen Figuren und deren Farben durchgeführt. Bei dieser Suchanfrage gibt es zwei Möglichkeiten. Entweder man sucht die Konturen, die in dem Bild vorhanden sein sollen. Oder man gibt zusätzlich auch noch die Farbe einzelnes Bereiches. Wir markieren Bereich und füllen ihn mit Farbe.

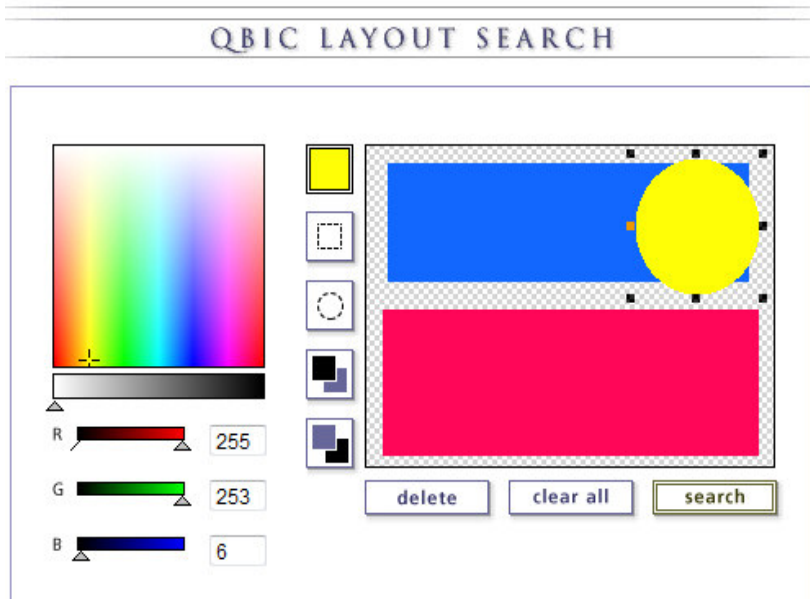


Abb3. Suche nach Konturen

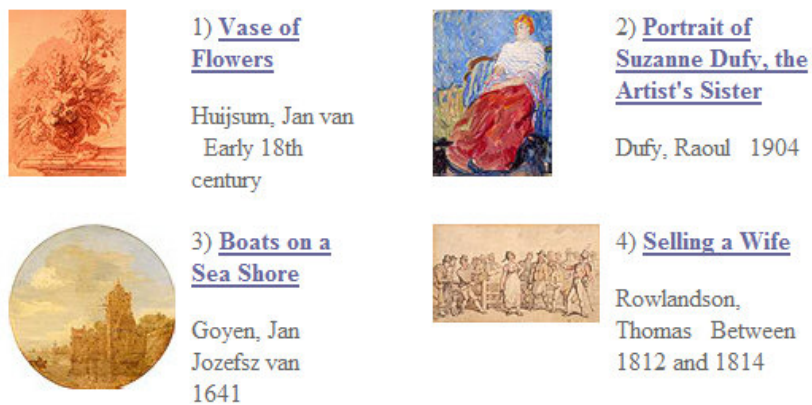


Abb. 4. Ergebnisliste der Suche nach farbigen Konturen

5. Fazit

Man kann schon behaupten, dass die inhaltsbasierte Multimediasuche eine zukunftsweisende Technologie ist und auch noch lange bleiben wird. Es sind bestimmt auch noch nicht alle Merkmale einer Bild- bzw. Videodatei untersucht worden sind. Und es werden eventuell noch weitere Suchkriterien bezüglich des Inhalts dazukommen.

Allerdings kann man nicht erwarten, dass die inhaltsbasierte Suche eine exakte Übereinstimmung von Suchkriterien und Bildmerkmalen liefert. Die Erfahrung hat gezeigt, dass in vielen Anwendungen die Suchergebnisse nicht den Erwartungen entsprechen. Häufig wurden als Ergebnis irrelevante Objekte zurückgeliefert. Das liegt an mehreren Problemen.

Erstens, das Extrahieren der Bildinformation, bzw. deren Berechnungen, sehr komplex ist. Man misst ja nur die Ähnlichkeit zwischen Suchkriterien und Bildinformation.

Zweitens, es existiert immer noch keine universelle Anfragesprache zur inhaltsbasierten Suche in den Datenbanken. Jedes Multimediadatenbanksystem benutzt die eigene Spezialsprache, bzw. Suchmethode.

Ebenso ein großes Problem bleibt auch die semantische Lücke. Die aus dem Bild extrahierte Digitalwerte können die menschliche Interpretation nicht nachbilden.

Deswegen werden diese Techniken bestimmt noch lange die herkömmliche Suche nicht ersetzen, sondern die nur ergänzen.

Quellen:

[wikipedia]http://de.wikipedia.org/wiki/Content_Based_Image_Retrieval

[Flickner] Flickner, M.; Query by Image and Video Content: The QBIC System

[Pittner] Doreen Pittner: Analyse kommerzieller ORDB-Bild-Retrieval-Systeme und Untersuchung notwendiger Erweiterungen zur Verbesserung der Retrieval-Qualität

[Schmitt 2005] Schmitt, I.: Ähnlichkeitssuche in Multimedia-Datenbanken: Retrieval, Suchalgorithmen und Anfragebearbeitung

[Kritzenberger] Kritzenberger, H. : Multimediale und interaktive Lernräume

[TemplMatch] Zorn, R. : Template Matching. Universität Paderborn

[ibm]<http://www.qbic.almaden.ibm.com/>

[herm]<http://www.hermitagemuseum.org/>