

Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Vortragsausarbeitung

Raoul Pascal Pein

Nutzung von Metainformation für die Navigation in (verteilten)
Datenbeständen

Zusammenfassung

Dieses Dokument dient dazu, den Inhalt des gleichnamigen Vortrages im Rahmen der Lehrveranstaltung *Anwendungen I* im Masterstudiengang *Verteilte Systeme* an der HAW Hamburg näher zu beschreiben.

Im Folgenden wird die Notwendigkeit für Suchwerkzeuge in großen digitalen Datenbeständen näher betrachtet. Es wird ein kurzer Überblick über die Probleme und mögliche Lösungswege bei der Realisierung dieser Werkzeuge gegeben. Die Definition von Ähnlichkeit, sowie die Extraktion dieser Informationen aus Dokumenten stellen die theoretische Grundlage dar, auf der eine maschinelle Suche aufsetzen kann. Der Vergleich bestehender Systeme zeigt, dass es noch viele Ansatzpunkte gibt, für die noch keine zufriedenstellenden Lösungen existieren.

Abschließend wird das Konzept des geplanten Projektes vorgestellt, das sich auf die Ergebnisse dieser Analyse stützt. Dieses Projekt ist in ein größeres Projekt im Bereich *Computer Supported Collaboration (CSC)* eingebettet.

Inhaltsverzeichnis

1 Einführung	3
2 Analyse derzeitiger Technologien	3
2.1 Ähnlichkeit	3
2.2 Erfassung der Aspekte/Metadaten	5
2.3 Navigation und Suchanfragen	5
2.4 Mehrdimensionale Suchräume	7
2.5 Bestehende Systeme	8
3 Projektbeschreibung	11

1 Einführung

Die Suche nach digitalen Informationen wird im Laufe der Zeit immer aufwendiger. Gerade durch Technologien wie das Internet ist es sehr einfach, vielen Menschen den Zugriff auf digitale Daten zu ermöglichen. Inzwischen ist das Angebot so groß, dass es schwierig ist, nur die relevanten Informationen aus einer Fülle von Quellen herauszusuchen.

Der Autor Hartmut Wittig beschreibt diese Situation folgendermaßen: „[...] users of information systems do not need more information, but information which better corresponds to their interests and needs, i.e. information of higher quality.“ [15]

Eine andere Sichtweise von Fernando Pereira und Rob Koenen ist sogar noch drastischer: „Content that cannot be easily found is like content that does not exist, [...]. The easier it becomes to produce content, the faster the amount of content grows and the more complex the problem of managing content gets.“ [2]

An dieser Stelle soll daher die Problematik hinter dem „Information Retrieval“ beleuchtet werden. Der Schwerpunkt liegt hierbei auf der Suche nach pixelbasierten Bildern. Hochwertige Suchmaschinen für Bilddokumente sind im Internet noch nicht zu finden. Bisher existieren größtenteils Prototypen oder Speziallösungen für bestimmte Anwendungsfelder, wie z.B. QBIC [8] oder Flickr [5].

Anders als Texte lassen sich Bilder nicht automatisch nach Schlagwörtern oder Phrasen indexieren. Bei großen Datenmengen ist eine manuelle Annotation, die die Nutzung von Suchstrings ermöglicht, sehr zeitraubend und kostspielig. Eine automatische Extraktion der Bildsemantik wäre wünschenswert. Dies ist aber mit dem heutigen Stand der Technik nicht oder nur für sehr eng begrenzte Bereiche möglich. Die Suche in großen Bildbeständen muss demnach auf andere Weise realisiert werden [4].

2 Analyse derzeitiger Technologien

2.1 Ähnlichkeit

Jede Suche hat das Ziel, aus einer nicht oder nur schwer überschaubaren Menge Informationen den relevanten Teil zu finden. Dies geschieht in der Regel mit Vorgaben, die das Gesuchte möglichst genau beschreiben. Diese Vorgaben werden mit dem Informationsbestand verglichen um Identität oder Ähnlichkeit festzustellen. Da pixelbasierte Bilder meistens viel unstrukturierte Information enthalten, ist eine Suche auf Identität (= identisches Pixelmuster) nur in Ausnahmefällen sinnvoll. Meistens interessieren den Suchenden möglichst ähnliche Bilder, wobei Details wie die Auflösung eine eher untergeordnete Rolle spielen.

Bei der Definition von Ähnlichkeit verschiedener Bilder gibt es viele Herangehensweisen. Jeder Mensch nimmt Bilder auf eine individuelle, subjektive Weise

Layer 1	primitives: color, texture, geometric shapes and spatial distribution
Layer 2	entities: object instances or individuals
Layer 3	abstract entities: activities, states, ...

Abbildung 1: 3-Schichten-Modell für Bilder [4]

wahr. Angefangen bei der Farbwahrnehmung, über die Erkennung bekannter Strukturen bis zu persönlichen Erinnerungen bei der Betrachtung eines Bildes können unterschiedlichste Kriterien herangezogen werden. Außerdem können Menschen selbst bei unterschiedlicher Beleuchtung oder variierenden Betrachtungswinkeln Motive wiedererkennen.

Weiterhin gibt es definierte, mathematische Verfahren, die in der Regel objektiv arbeiten. Beispiele sind die Extraktion von Histogrammen oder Konturen aus einem Bild. Diese sollten möglichst unempfindlich gegenüber Unschärfe, Kratzer, Beleuchtung oder andere Störungen sein. Liegen dem Suchprogramm diese Informationen vor, kann es über die Berechnung einer mathematischen Distanz die Ähnlichkeit zweier Bilder ermitteln. Eine effiziente Implementierung sollte hierbei möglichst kompakte, aussagekräftige und schnell vergleichbare Indexdaten generieren. Diese Verfahren lassen sich meist vollständig automatisieren.

Eine sehr mächtige Methode ist der Vergleich über die Zugehörigkeit zu bestimmten Kategorien oder die Semantik eines Bildes. Ist genau bekannt, was auf zwei Bildern dargestellt ist, ist die Bestimmung der Ähnlichkeit relativ einfach und naheliegend. Leider ist es nicht trivial, diese Daten automatisch aus einem Bild zu extrahieren.

Eine Unterscheidung verschiedener Ebenen des Bildinhaltes wird im 3-Schichten-Modell (Abb. 1) vorgestellt [4]. Schicht 1 behandelt die grundlegenden Primitive eines Bildes. Hier werden Farben, Texturen, geometrische Formen und die räumliche Verteilung von nicht näher spezifizierten Objekten beschrieben (z.B. „grün, unregelmäßig“). Schicht 2 erfordert bereits eine gewisse Sachkenntnis, da hier mit konkreten Entitäten aus der wirklichen Welt gearbeitet wird. Die Beschreibung umfasst bestimmte Instanzen oder Individuen, die abgebildet sind (z.B. „ein Auto“). Schicht 3 ist die höchste Abstraktionsschicht von den zu Grunde liegenden Pixeln. Sie bildet Aktivitäten oder Zustände ab, die nicht nur aus dem Vorhandensein verschiedener Elemente abhängt (z.B. Emotionen).

Ein Beispiel für ein konkretes Bild:

Ebene 1: Die obere Hälfte ist hellblau mit weißen Anteilen, die untere Hälfte grün. Verschiedenfarbige Flächen sind über das Bild verteilt.

Ebene 2: Die obere Hälfte ist Himmel, die untere Hälfte ist Rasen. Darauf stehen mehrere Personen.

Ebene 3: Das Bild stellt ein Handballspiel auf einem Sportplatz dar.

2.2 Erfassung der Aspekte/Metadaten

Bevor ein Retrievalsystem effizient arbeiten kann, muss ein Index über alle verwalteten Daten angelegt werden. Je größer der Datenbestand ist, desto wichtiger ist es, diesen Vorgang zu automatisieren. Andernfalls muss diese Arbeit von Menschen übernommen werden, was ungleich länger dauert.

Für die automatisierte Erfassung spricht weiterhin, dass die angewandten Algorithmen objektiv arbeiten und für ein Bild/Dokument immer die selben Ergebnisse liefern. Der Hauptnachteil liegt in dem Problem der mangelnden Intelligenz und Flexibilität der Software. Mit ihr ist es bisher nicht möglich, viel mehr Information als nur einfache Konstrukte zu gewinnen. Aus Bildern können lediglich Primitive (Ebene 1) extrahiert werden. Abstraktere Informationen bleiben dem Algorithmus in aller Regel verborgen.

Der gegenteilige Ansatz mit manueller Erfassung erfordert einen ungleich höheren Arbeitsaufwand. Jedes Dokument muss einzeln von einem Administrator gesichtet und beschrieben werden. Dabei kann es vorkommen, dass subjektive oder gar widersprüchliche Daten in das System gelangen. Der große Vorteil bei diesem Ansatz liegt in der Möglichkeit, gerade abstrakte Merkmale zu erfassen. Werden Spezialisten für das gegebene Fachgebiet eingesetzt, ist eine sehr hohe Qualität des Index erreichbar. Menschen können auch in Ausnahmesituationen flexibel reagieren, gegebenenfalls nachrecherchieren und Zusammenhänge erkennen.

Als optimale Lösung kann sich ein halbautomatischer Ansatz herausstellen. Die Erfassung von Primitiven kann vollständig automatisiert ablaufen. Nur die abstrakten Eigenschaften müssen bei Bedarf von Hand nachgetragen werden. Eine automatisierte Klassifikation von Bildern ist in [11] beschrieben. Dem Algorithmus werden im Voraus repräsentative Beispiele für die einzelnen Klassen vorgegeben. Hinterher sollten die Ergebnisse aber noch einmal gesichtet und von Hand korrigiert werden.

2.3 Navigation und Suchanfragen

Für die Navigation in großen, unstrukturierten Datenbeständen existieren verschiedene Möglichkeiten, die gesuchte Information zu finden.

Verlinkung Ein Beispiel für einen Datenbestand, der zum größten Teil aus Textdokumenten besteht, ist das World Wide Web. Das Standardformat für diese *Hyper-*

text-Dokumente ist die Seitenbeschreibungssprache HTML. Diese erlaubt es, über Hyperlinks direkt auf andere Dokumente zuzugreifen. Eine erweiterte Variante, die ihren Schwerpunkt auf Multimediadokumente legt, wird auch *Hypermedia* bezeichnet. Diese Dokumente erlauben es, an vordefinierte Stellen innerhalb von Audio- oder Videodateien zu springen. Weiterhin wird auch die Synchronisation zwischen verschiedenen Medien darüber realisiert.

Diese Verlinkung verschiedener Dokumente erlaubt eine gezielte Navigation. Der Benutzer kann sich an vorhandenen Strukturen „entlanghangeln“, um an die gewünschte Information zu kommen. Damit diese Navigation funktioniert, müssen die Daten stimmig aufbereitet werden. Dies geschieht im Endeffekt immer manuell, da bisher nur Menschen die sinnvollen Verbindungen zwischen Dokumenten erkennen können. Unsinnige oder fehlende Links führen schnell in die falsche Richtung oder Sackgassen. Dadurch kommt der Nutzer gar nicht oder nur über Umwege zu seinem Zieldokument.

Suche Eine Suche macht immer dann Sinn, wenn die vorliegenden Informationen nicht ausreichend miteinander verknüpft sind und der Datenbestand zu groß ist, um diese Verknüpfungen mit vertretbarem Aufwand zu realisieren.

Die einfachste Variante der sequentiellen Suche ist ab einer gewissen Datenmenge praktisch nicht mehr anwendbar. Daher muss dem Betrachter ein möglichst kleiner Ausschnitt der Daten präsentiert werden, der genau die gesuchten Informationen enthält. Dabei ist es wichtig, gleichzeitig möglichst viele der irrelevanten Datensätze auszufiltern. Die resultierende Ergebnismenge kann dann manuell nach relevanten Informationen durchsucht werden.

Die Qualität der Ergebnisse steht und fällt mit der Genauigkeit der Suchanfrage und den Möglichkeiten der Suchmaschine. Oft wird hier mit dem Begriff der Ähnlichkeit gearbeitet, da die Anfrage selten exakt mit dem Ergebnis übereinstimmt. Dadurch entsteht bei der Abgrenzung zwischen *Treffer* und *irrelevant* eine gewisse Unschärfe.

Mögliche Anfragen an ein Retrieval-System Bei der Formulierung einer Suchanfrage stehen dem Nutzer generell 3 verschiedene Kriterientypen zur Verfügung.

Zur Verkleinerung des Suchraums sind *Filter* nützlich. Mit ihrer Hilfe kann man beispielsweise die Ergebnisse auf einem bestimmten Zeitraum (Erstellung des Dokumentes, abgebildeter Zeitpunkt) oder eine bestimmte Kategorie begrenzen. Für gewöhnlich arbeiten Filter mit einer harten Trennung zwischen relevant und irrelevant. Dadurch kann eine Suche über entsprechend aufgebaute Indexe in der Regel sehr schnell erfolgen [12]. Die Bildung der zwei Teilmengen kann auch im Bedarfsfall mit Fuzzy-Logik geschehen, so dass auch Ergebnisse in Randbereichen eine gewisse Relevanz erhalten [15].

Für das „Content Based Image-Retrieval“ ausschlaggebend ist die Suchanfrage mit einem *Beispiel*. Hierbei wird dem Programm ein Bild präsentiert, für das ähnliche Bilder in der Sammlung gefunden werden sollen (vgl. Abschnitt 2.1). Bei Textdateien

kann dieses Verfahren ebenfalls helfen, um vergleichbare Dokumente aufzuspüren. Diese Anfragen können generell nicht exakt formuliert werden, da der Anwender in diesem Fall oft selber nicht so genau weiß, was er haben möchte. Aus diesen Anfragen kann der Computer automatisch gewisse Eigenschaften (z.B. Histogramme) extrahieren die helfen, den Vergleich durchzuführen. Eine scharfe Abgrenzung von Treffern zu uninteressanten Dokumenten ist hierbei meist unsinnig. Für die Ergebnisse können Ranking-Werte erstellt werden, die die Ähnlichkeit zur Anfrage beispielsweise in Prozent angeben. Eine effiziente Indexierung kann bei komplexen Eigenschaften sehr aufwendig werden. Hier sind beispielsweise mehrdimensionale Strukturen wie R-Bäume [6] notwendig.

Eine dritte Kategorie bilden die *Metadaten* eines Dokumentes. Ob diese eine exakte Klassifizierung oder auch Unschärfe bedeuten, ist abhängig von dem jeweiligen Informationstyp. Auf alle Fälle müssen diese Informationen explizit bei der Suchanfrage eingegeben werden, da sie nicht direkt aus einem Beispieldokument extrahiert werden können. Beispiele hierfür sind Schlüsselwörter oder die Semantik eines Dokumentes. Diese werden in der Regel von einem Experten definiert. Anders als gewöhnliche Filter unterstützen diese Anfragen auch die Generierung von Ranking-Werten (z.B. um Tippfehler, verschiedene Schreibweisen oder ähnliche Semantik zu erfassen). Für diesen Zweck bieten sich Techniken aus dem Bereich Semantic Web und Topic Maps [3] an.

2.4 Mehrdimensionale Suchräume

Wie bereits in Abschnitt 2.1 erwähnt, existieren diverse Aspekte, die alle einen bestimmten Blickwinkel auf ein Dokument darstellen. In Abhängigkeit von den Bedürfnissen des Anwenders erscheint es zweckmäßig, diesem bei der Anfrageformulierung die Wahl zwischen verschiedenen Aspekten zu geben. Möglicherweise ist er an bestimmten Begriffen (Suchstring) interessiert, die mit dem Dokument verküpft sind. Vielleicht wird auch gerade ein Bild benötigt, das bestimmte Farben (Histogramm) aufweist. In beiden Fällen sieht die Anfrage grundlegend verschieden aus. Während ersteres mit einem einfachen String bewerkstelligt werden kann, ist es ungleich schwieriger, die gewünschten Farben eines Bildes anzugeben.

Betrachtet man das aktuelle Angebot an Image Retrieval Systemen im Internet, wird deutlich, dass viele Systeme lediglich Suchstrings anbieten. Echtes CBIR hingegen beschränkt sich oft nur einen Aspekt, nach dem die Ähnlichkeit bestimmt wird [14]. Ein aus der Masse herausragendes System ist MARVel [13], das die Auswahl diverser verschiedener Aspekte ermöglicht.

Ein System das die Nutzung verschiedener Aspekte/Blickwinkel anbietet, muss in der Lage sein, einen mehrdimensionalen Suchraum zu verwenden. Jeder Aspekt stellt hierbei eine Dimension dar, die mit den anderen Dimensionen in Einklang gebracht werden muss. Das Problem wird für 2 Aspekte in Abb. 2 visualisiert. Der Schnittpunkt der beiden Linien stellt das optimale Ergebnis dar. Die heller werdenden Kreise repräsentieren das sich verschlechternde Ranking für die Objekte mit

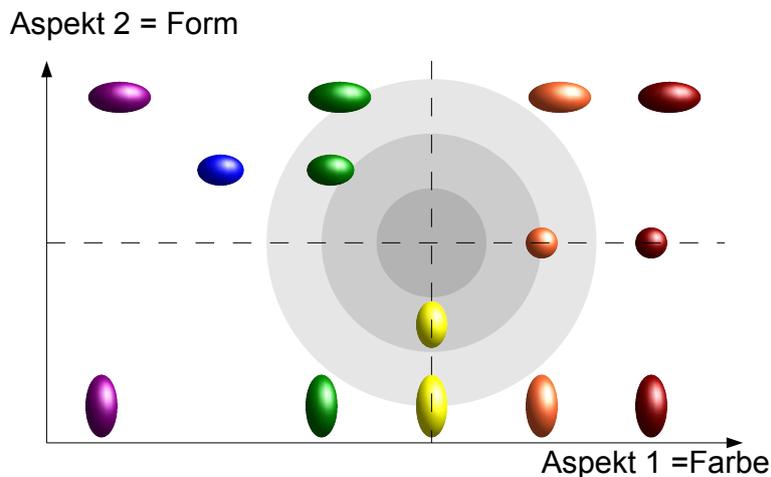


Abbildung 2: Mehrdimensionale Suchräume - Suche nach einer gelben Kugel

wachsender Entfernung zur Anfrage.

Als einfaches Verfahren bietet es sich an, die verschiedenen Dimensionen auf ein gemeinsames Maß herunterzubrechen. Jeder Ähnlichkeitswert kann und muss auf ein eindimensionales Ranking abgebildet werden (0%-100% Ähnlichkeit). Über die resultierenden Teilrankings jedes Aspektes kann ein gewichteter Mittelwert gebildet werden. So erhält jedes Dokument eine einfache Zahl als Ranking, mit der die Suchergebnisse sortiert dargestellt werden können.

In diesem Fall gilt:

$$r_x = \frac{1}{\sum_{f=1}^n w^f} * \sum_{f=1}^n w^f * r_x^f$$

wobei r_x das zusammengefasste Ranking für Dokument x ist, w^f das relative Gewicht für einen Aspekt (feature) und r_x^f das berechnete Teilranking eines Dokumentes für Aspekt f . [10]

2.5 Bestehende Systeme

Dieser Abschnitt stellt aktuelle Technologien und Standards vor, die im Zusammenhang mit Retrieval Systemen stehen.

Beschreibungsstandards Für die textuelle Annotation haben sich mehrere Sprachen etabliert. Als Grundlage dient in der Regel die *Standard Generalized Markup Language* (SGML), die eine individuelle Strukturierung textbasierter Dokumente erlaubt. Eine vereinfachte Teilmenge dieser Sprache ist die *Extensible Markup Language* (XML). Mit ihrer Hilfe können selbst komplexe Zusammenhänge flexibel modelliert werden. Für die Repräsentation von Metadaten im Internet kann das Re-

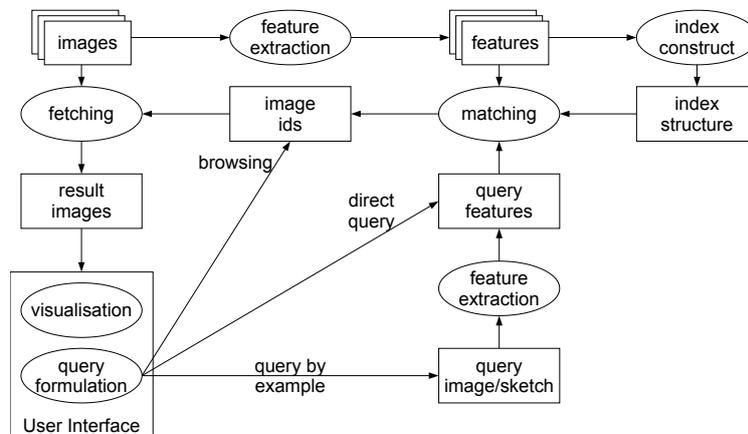


Abbildung 3: Standard CBIR-Architektur [14]

source Description Framework (RDF) verwendet werden. Dieses erlaubt Beschreibungen wahlweise als grafisches Modell oder auf Basis von XML.

Zur Repräsentation von Multimediadaten wurde das *Multimedia Content Description Interface* von der *Moving Picture Experts Group* entworfen (MPEG-7). Diese Sprache ist speziell auf die Eigenschaften von Videos und Audiodateien zugeschnitten [2].

Architektur Abbildung 3 zeigt eine Standardarchitektur für „Content Based Image Retrieval“, wie sie von Veltkamp [14] beschrieben wird. Im Kern wird aus einem Satz Bildern ein Index über die extrahierten Aspekte (features) jedes Bildes generiert. Suchanfragen werden entweder direkt oder via Beispielbild gestellt. Über einen Matchingalgorithmus wird die Anfrage mit dem bestehenden Index verglichen und dem Benutzer eine Auswahl an Treffern präsentiert.

CBIR-Prototyp Ein Beispiel für ein einfaches CBIR-System ist der Prototyp, der im Zuge einer Diplomarbeit [9] entstanden ist (Abb. 4). Dieses System erlaubt Anfragen über Beispiele und einen Suchstring. Der Nutzer kann bestimmen, welche Aspekte (Histogramm, ortsabhängige Histogramme, Schlagwörter) bei der Suche berücksichtigt werden sollen. Zusätzlich ist es ihm möglich, die relative Gewichtung der einzelnen Aspekte zueinander auf seine Bedürfnisse einzustellen.

UIMA-Architektur (IBM Research) Die Firma IBM arbeitet an einer offenen Architektur für Informationsmanagement, genannt „Unstructured Information Management Architecture“ (UIMA) [7]. Die Idee ist die Schaffung einer nutzbaren, skalierbaren und erweiterbaren Plattform, die die Entwicklung von Informationssystemen vereinfacht.

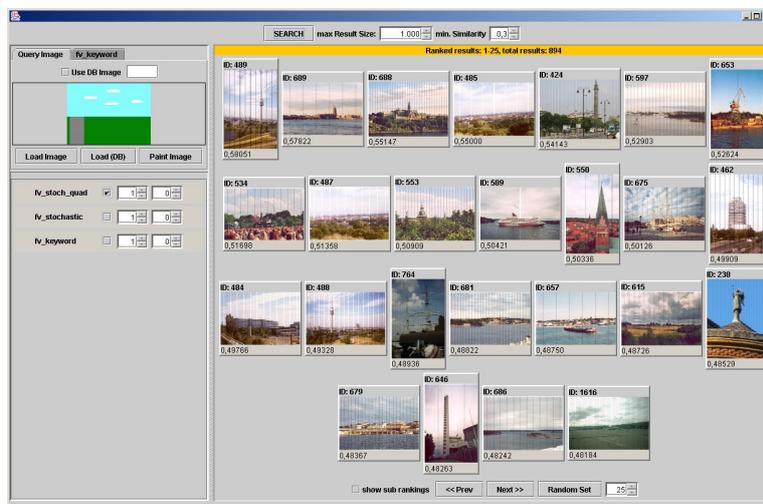


Abbildung 4: Inhaltsbasierte Bildersuche [9]

Das Framework ist in Java realisiert und wird auf open source Basis angeboten. Die Architektur 2.5 hat deutliche Parallelen zu der Architektur von Veltkamp [14]. Der wesentliche Unterschied besteht in der zusätzlichen Unterstützung verschiedener Datentypen und einer semantischen Suche.

Relevance Feedback Eine gewöhnliche Suchmaschine liefert bei wiederholter identischer Anfrage identische Ergebnisse. Nur nach einer Änderung des Indexes durch neue oder veränderte Informationen kann das Ergebnis variieren.

Wenn eine Suchmaschine lernt, welche Ergebnisse brauchbar sind und welche nicht, kann auf diese Weise das Suchergebnis im Laufe der Zeit verbessert werden. Nach jeder Suche könnte der Nutzer eine Möglichkeit erhalten, das Ergebnis zu bewerten. Innerhalb einer Sitzung könnte man so schrittweise die Relevanz der gefundenen Einträge erhöhen.

Ist die Suchmaschine in der Lage, diese Informationen über mehrere Sitzungen hinweg zu speichern, resultiert dies in einer Anpassung der Ergebnisse an die Bedürfnisse des vorherigen Nutzers. Es ist also denkbar, individuelle Benutzerprofile für die Suchmaschine anzulegen, die maßgeschneiderte Ergebnisse liefern.

Weiterhin kann auf diese Weise das Angebot für alle Nutzer verbessert werden. Allerdings ist dabei darauf zu achten, dass das Interesse verschiedener Nutzer stark variieren kann. Es sollte nur Feedback erlaubt sein, das bestimmte Informationen nicht komplett ausblendet oder zu falschen Ergebnissen führt.

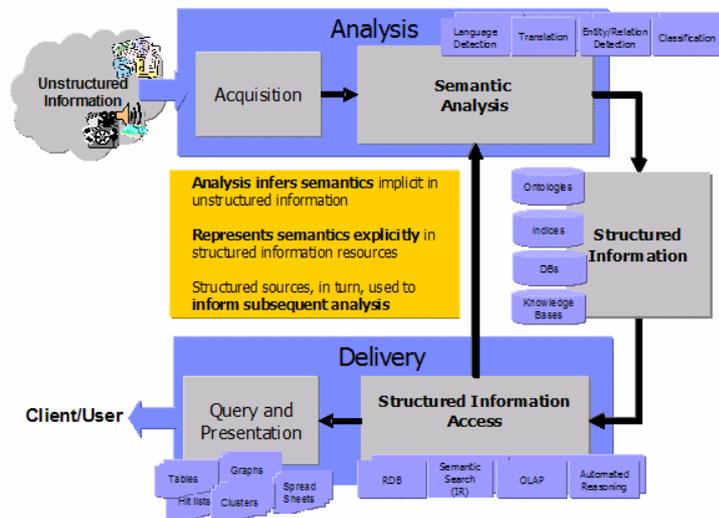


Abbildung 5: UIMA-Architektur [7]

3 Projektbeschreibung

Das Hauptziel des Projektes ist die Entwicklung eines Retrievaldienstes mit einer offenen Schnittstelle (z.B. über Webservices). Diese soll das Anfordern einer Liste mit ähnlichen Bildern für ein Bild ermöglichen, die von anderen Applikationen verwendet werden kann.

Als Grundlage für dieses Programm dient der Prototyp, der bereits in der Diplomarbeit entwickelt wurde.

Zusätzlich zu dem Kerndienst für das gemeinsame Collaborative Workspace Projekt sollen Erweiterungen für eine effiziente Bilder-/Dokumentensuche integriert werden. Hierzu gehören Technologien zur Erfassung des Dokumenteninhalts auf verschiedenen Ebenen (z.B. Schlagwörter, Konturen, Histogramme, Semantik) sowie ein Relevance Feedback zur iterativen Verbesserung der Ergebnisqualität.

Die Annotation der Bilder soll spontan von beliebigen Anwendern der Software geschehen können. Dadurch wird die Arbeitslast auf viele Schultern verteilt. Als Vorbild dient hierbei das Wiki-Prinzip (z.B. MediaWiki [1]). Als Grundlage für die Annotation sollen nach Möglichkeit etablierte Standards, wie beispielsweise MPEG-7 [2] genutzt werden.

Außerdem soll untersucht werden, in wie weit eine verteilte Suche über stationäre und mobile Geräte realisiert werden kann. Bei jeder Suche kann ein Client alle verfügbaren Dienste auffindig machen und dort Anfragen starten. Die Ergebnisse werden dann zusammengefasst und dem Benutzer präsentiert. Wo das Originalbild liegt (Server, Notebook, PDA, ...) und wie es angefordert werden muss, wird von der Software automatisch geregelt.

Abschließend wird ein Beispiel für ein denkbare Szenario gegeben, in dem

das beschriebene Projekt helfen kann.

Szenario Eine Werbefirma benötigt für eine Collage auf einem Plakat eine Auswahl an passenden Bildern. Da die Bilddatenbank eine unüberschaubare Größe besitzt, wird eine Image Retrieval Software eingesetzt. Diese muss nun eine entsprechende Suchanfrage erhalten, die alle relevanten Informationen zu den gesuchten Bildern enthält. Als Kriterien werden genannt:

Zeitraum 1990-2006 (altmodisch wirkende Bilder passen bei der Werbekampagne nicht)

Beispielbild hellblauer Himmel, blaugrünes Meer, weißer Strand

Thema Seereise

Diese Anfrage besteht aus drei verschiedenen Elementen. Die Angabe des Zeitraumes schränkt hierbei den Suchraum erheblich ein. Das Beispielbild definiert eine grobe Vorstellung davon, wie das benötigte Bild ungefähr auszusehen hat. Eine Angabe des Themengebietes dient dazu, möglichst nur Ergebnisse zu präsentieren, die inhaltlich passen. Die resultierende Ergebnismenge sollte dann einen sehr kleinen Ausschnitt aus dem Datenbestand enthalten, der gleichzeitig eine hohe Ähnlichkeit mit der Anfrage hat.

Literatur

- [1] Mediawiki. Available from: <http://www.mediawiki.org/>.
- [2] Thomas Sikora B. S. Manjunath, Philippe Salembier, editor. *Introduction to MPEG-7 - Multimedia Content Description Interface*. John Wiley & Sons Ltd., 2002. B. S. Manjunath: University of California, Santa Barbara, USA; Philippe Salembier: Universitat Politècnica de Catalunya, Barcelona, Spain; ISBN 0-471-48678-7.
- [3] Andreas Christensen. Semantische Anreicherung von Suchanfragen auf Basis von Topic Maps. Diplomarbeit.
- [4] J.P. Eakins and M.E. Graham. Content-based Image Retrieval. A Report to the JISC Technology Applications Programme. Technical report, University of Northumbria at Newcastle, January 1999. Available from: <http://www.unn.ac.uk/iidr/VISOR>.
- [5] Flickr - Photo Sharing, 2006. Available from: <http://www.flickr.com/>.
- [6] A. Guttman. R-trees: a dynamic index structure for spatial searching. In *ACM SIGMOD Conf. on the Management of Data*, pages 47–57, 1984.

- [7] IBM. Unstructured Information Management Architecture. Available from: <http://www.research.ibm.com/UIMA/>.
- [8] W. Niblack, X. Zhu, J. Hafner, T. Breuel, D. Ponceleón, D. Petkovic, M. Flickner, E. Upfal, S. Nin, S. Sull, B. Dom, B.-L. Yeo, S. Srinivasan, D. Zivkovic, and M. Penner. Updates to the QBIC system. *Retrieval for Image and Video Databases VI*, 3312:150–161, 1998.
- [9] Raoul Pascal Pein. Multi-Modal Image Retrieval - A Feasibility Study. Diplomarbeit.
- [10] Raoul Pascal Pein and Zhongyu (Joan) Lu. Content Based Image Retrieval by Combining Features and Query-By-Sketch. In *The 2006 International Conference on Information & Knowledge Engineering*, 2006.
- [11] Guoping Qiu, Xia Feng, and Jianzhong Fang. Compressing histogram representations for automatic colour photo categorization. *Pattern Recognition*, 37:2177–2193, 2004.
- [12] Raghu Ramakrishnan. *Database Management System*. The McGraw-Hill Companies, Inc., 1998.
- [13] John R. Smith. Marvel: Multimedia analysis and retrieval system. Whitepaper, Intelligent Information Management Dept. IBM T. J. Watson Research Center, 2005.
- [14] Remco C. Veltkamp and Mirela Tanase. Content-Based Image Retrieval Systems: A Survey. Technical Report UU-CS-2000-34, Department of Computing Science, Utrecht University, October 2002.
- [15] Hartmut Wittig. *Intelligent Media Agents: Key technology for Interactive Television, Multimedia and Internet Applications*. GWV-Vieweg, 1999.