

# Zweitveröffentlichung



Henrich, Andreas

## Multimedia Information Retrieval

Datum der Zweitveröffentlichung: 20.11.2023

Verlagsversion (Version of Record), Zeitschriftenartikel

Persistenter Identifikator: urn:nbn:de:bvb:473-irb-918714

### Erstveröffentlichung

Henrich, Andreas (2009): „Multimedia Information Retrieval“. In: Information technology : methods and applications of informatics and information technology, Jg. 51, Nr. 6, S. 336-342, Berlin: De Gruyter Oldenbourg, doi: 10.1524/itit.2009.0560.

### Rechtehinweis

Dieses Werk ist durch das Urheberrecht und/oder die Angabe einer Lizenz geschützt. Es steht Ihnen frei, dieses Werk auf jede Art und Weise zu nutzen, die durch die für Sie geltende Gesetzgebung zum Urheberrecht und/oder durch die Lizenz erlaubt ist. Für andere Verwendungszwecke müssen Sie die Erlaubnis des/der Rechteinhaber(s) einholen.

Für dieses Dokument gilt das deutsche Urheberrecht.

# Multimedia Information Retrieval

## Multimedia Information Retrieval

Andreas Henrich, Otto-Friedrich-Universität Bamberg

**Zusammenfassung** Der Umfang, in dem multimediale Daten verfügbar sind, nimmt ständig zu. Dies gilt für das Internet ebenso wie für die Daten auf dem heimischen PC oder auch auf einem PDA. Eine wesentliche Herausforderung stellt dabei die Suche nach Medienobjekten dar. Der vorliegende Artikel geht auf grundlegende Konzepte und Problemstellungen zur Suche nach Medienobjekten und multimedialen Dokumenten ein und gibt einen Überblick über die für verschiedene Medien-

typen existierenden Ansätze. ▶▶▶ **Summary** The extent to which multimedia data are available is steadily increasing. This is true for the Internet as well as for the data on a local PC or on a PDA. A major challenge in this respect is the search for media objects. The present article deals with fundamental concepts and issues of search for media objects and multimedia documents, and gives an overview of the approaches existing for the different media types.

**Schlagwörter** Suche, Multimedia, Suchmaschine, Bildsuche ▶▶▶ **Keywords** H.3.3 [Information Systems: Information Storage and Retrieval: Information Search and Retrieval]; I.7 [Computing Methodologies: Document and Text Processing]; Search, multimedia, search engine, image retrieval

### 1 Einleitung

Mit der Flut von multimedialen Daten im Internet, im Intranet und auf persönlichen Rechnern und Endgeräten wird die effiziente Verwaltung multimedialer Daten immer wichtiger. Hierzu gehört insbesondere eine leistungsfähige Suche. Die Aufgabenstellungen reichen von der Bildsuche auf mobilen Endgeräten über die Suche nach ähnlichen Bewegungssequenzen beim Motion-Tracking bis zur Suche in 3D-Daten.

Im vorliegenden Artikel wird ein Überblick über die verschiedenen Gebiete gegeben. Dabei werden wir uns nach einführenden Überlegungen mit der Suche nach einzelnen Medienobjekten beschäftigen (Text, Bilder, Audiodaten, Videos...). Zunehmende Bedeutung gewinnt aber auch die Suche nach und in multimedialen Dokumenten, in denen Medienobjekte zu einem strukturierten multimedialen Dokument zusammengefasst sind. Hierauf werden wir in Abschnitt 7 eingehen. Schließlich werden wir in Abschnitt 8 aktuelle Trends ansprechen.

#### 1.1 Arten von Suchproblemen

Wenn wir von der „Suche“ nach Medienobjekten sprechen, dann können sich dahinter sehr verschiedene Aufgabenstellungen verbergen.

Ein Fall ist die gezielte Suche nach Medienobjekten, die wir bereits kennen und nun wieder finden möchten.

Hier werden Suchanfragen oft sehr kreativ formuliert. So wird z. B. nach einem Bild gesucht, bei dem links im Hintergrund ein blaues Auto stand.

Anders verhält es sich bei der Suche nach uns unbekanntem Medienobjekten, die wir nur inhaltlich beschreiben können. Aus der Websuche kennen wir hier die Formulierung unseres Informationswunsches mit wenigen Begriffen. Eine typische Suchmaschine sucht nach den Textdokumenten (Webseiten), die diese Begriffe enthalten *und* populär sind – wobei die Popularität in der Regel durch die Anzahl der auf die Webseite verweisenden Links, gewichtet mit der Popularität der jeweiligen Ausgangsseite, abgeschätzt wird.

Das beschriebene Modell kann auf andere Medientypen übertragen werden, wenn diese mit Text in Beziehung gesetzt werden können. Dies gilt für Bilder, die in Webseiten eingebettet sind, ebenso wie für Audiodateien, zu denen Metadaten oder Liedtexte verfügbar sind.

Sind zu einem Medienobjekt keine textuellen Beschreibungen vorhanden, so müssten die das Medienobjekt beschreibenden Begriffe zunächst automatisch bestimmt werden. Dabei stößt man auf die „semantische Lücke“. Die Bedeutung dieser Lücke erschließt sich am besten im Vergleich zum Medientyp *Text*. Ein Text setzt sich aus Wörtern zusammen, deren Grenzen durch Leer- und Satzzeichen leicht zu ermitteln sind. Die einzelnen Wörter

verraten viel über die in einem Text behandelten Themen. Betrachtet man dagegen ein Bild, so ist dieses aus Bildpunkten zusammengesetzt, aus denen letztlich Objekte gebildet werden. Allerdings sind weder die Grenzen der Objekte leicht und eindeutig zu ermitteln, noch kann die Bedeutung der gegebenenfalls segmentierten Pixel-Gruppen leicht erschlossen werden. Für eine begrenzte Domäne – z. B. Landschaftsaufnahmen – kann man zwar heute bereits ermutigende Erkennungsraten erzielen [19], eine allgemeine Lösung der Erkennung von Objekten in Bildern wird es aber wohl auf absehbare Zeit nicht geben.

Aufgrund der skizzierten Probleme hat sich das Paradigma des *Query-by-Example* (QbE) in verschiedenen Domänen etabliert. Dabei wird – im Fall der Bildsuche – ein Beispielbild vorgegeben, zu dem ähnliche Bilder gesucht werden. Die Ähnlichkeit bezieht sich allerdings in den meisten Fällen auf Low-Level-Features wie Farbverteilungen oder Textureigenschaften.

**1.2 Der Werkzeugkasten**

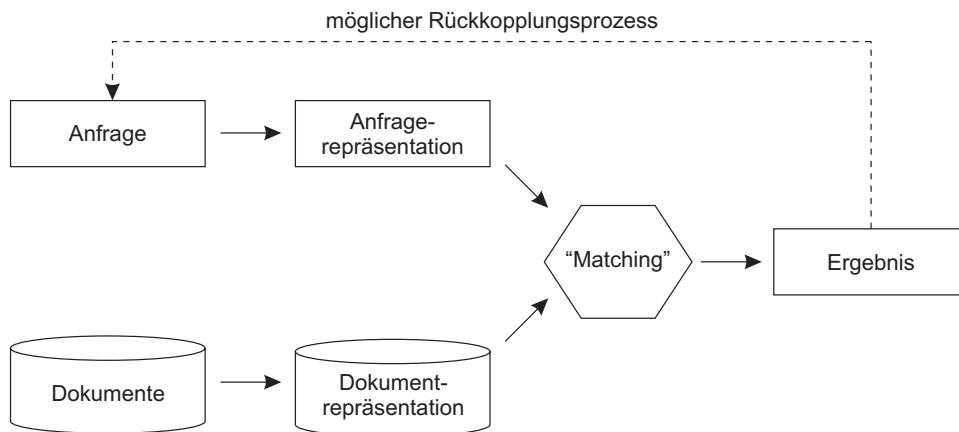
Die Suche nach Medienobjekten läuft in der Regel nach einem einfachen Grundschema ab (siehe Bild 1). Dabei gehen wir von einer QbE-Anfrage aus. In diesem Fall werden für die zu durchsuchenden Dokumente und für das Anfragedokument entsprechende Repräsentationen bestimmt. Bei Textdokumenten können dies hochdimensionale Beschreibungsvektoren sein, deren Komponenten die Relevanz für bestimmte Themen quantifizieren. Hat

man für die Anfrage und für alle Dokumente entsprechende Repräsentationen, so kann man unter allen Dokumenten diejenigen ermitteln, die zur Anfragerrepräsentation am ähnlichsten sind („*Matching*“). Der in der Abbildung angedeutete Rückkopplungspfad dient gegebenenfalls der Präzisierung der Anfrage („*Relevance Feedback*“).

Für verschiedene Medientypen existieren unterschiedliche Repräsentationen. Diese können sich bei Bildern z. B. auf Farbhäufigkeiten und bei Audiodateien auf Intensitäten in Frequenzbändern beziehen (Spalte *Feature-Extraktion* in Tabelle 1).

Betrachtet man Eigenschaften, die sich auf das gesamte Medienobjekt beziehen, spricht man von *globalen* Eigenschaften. So stellt ein Histogramm, das die Farbverteilung in einem Bild repräsentiert, eine globale Betrachtung dar. Bei einer *lokalen* Betrachtung wird nur ein Ausschnitt – ein Segment – des Medienobjekts betrachtet. Ein Beispiel wäre ein Paragraph in einem Text oder eine Szene in einem Film – die allerdings in der Regel aus den Inhalten der Frames erkannt werden muss. Auch bei Bildern müssen die einzelnen Segmente mit entsprechenden Verfahren erkannt werden [12]. Bei der Bildsuche kann man dann nicht nur nach Bildern suchen, die global ähnlich zu einem Anfragebild sind, sondern auch nach solchen, die ein passendes Segment enthalten.

Schließlich kann man versuchen durch den Übergang zu einem anderen Medientyp eine semantisch höher lie-



**Bild 1** Typische Architektur zur Suche nach Dokumenten oder Medienobjekten.

**Tabelle 1** Beispieloperationen für verschiedene Medientypen.

	Feature-Extraktion	Segmentierung	Konvertierung
Textdaten	Statistiken zu Wortvorkommen, ...	Gliederung in Kapitel, Abschnitte, Sätze, ...	als Bild drucken, vorlesen, ...
Bilddaten	Farbhistogramme, Textur, ...	regionenbasierte oder kantenbasierte Segmentierung, ...	Optical Character Recognition (OCR), ...
Audiodaten	Lautstärke, Rhythmus, Erkennung von Tönen, ...	Erkennung von Pausen, ...	Spracherkennung, ...
Videodaten	Bewegungserkennung, Shotlängen, ...	Szenenerkennung, ...	Extraktion von Schlüsselbildern, Erkennen von Untertiteln, ...

gende Ebene für die Suche zu nutzen. So kann man in Bildern mit OCR-Verfahren nach Texten suchen oder bei Videos auf die Untertitel zurückgreifen (Spalte *Konvertierung* in Tabelle 1).

## 2 Textsuche

Die Suche nach Textdokumenten [2] ist für Multimedia Information Retrieval von großer Bedeutung, weil sich andere Medienobjekte zum Teil in Text überführen lassen und Text in strukturierten multimedialen Dokumenten oft ein vielversprechender Ansatzpunkt für die Suche ist.

Einfach gesprochen werden Texte dabei im Sinne des Bildes 1 als Vektoren repräsentiert, bei denen die Komponenten die Vorkommenshäufigkeiten von Wörtern enthalten. Um hier zu guten Ergebnissen zu kommen, können verschiedene Verfahren vorgeschaltet werden. Als wesentliche Bereiche sind die Stopworteliminierung, die Grund- und Stammformreduktion, die Mehrwortgruppenbehandlung und die Betrachtung von Begriffsbeziehungen (Synonyme, zu spezielle Begriffe,...) zu nennen [8].

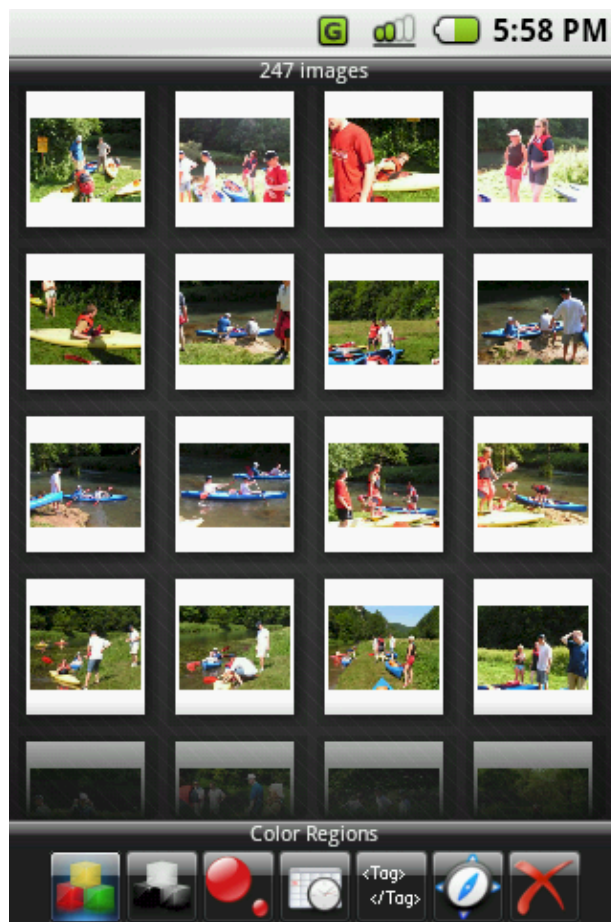
Zur Berechnung der Vektorkomponenten werden dann neben der Vorkommenshäufigkeit des jeweiligen Begriffs im Dokument auch seine Trennschärfe im Hinblick auf die ganze Kollektion und die Dokumentlänge als Normierungskriterium betrachtet. Als Beispiel für die konkrete Berechnung sei auf *BM25* (vgl. [16]) verwiesen.

## 3 Bildsuche

Bei der Bildsuche [4] ist zwischen einer Suche auf Basis von Metadaten (z. B. Exif-Daten oder Tags) und einer Suche auf Basis visueller Bildeigenschaften zu differenzieren. Während man sich bei einer Suche auf Basis von Tags im Bereich der Textsuche befindet, müssen im letzteren Fall Farb-, Textur- oder Formeigenschaften adressiert werden [9]. Eine besondere Rolle spielt die Invarianz der Eigenschaften. So sollten die in der Anfragebearbeitung adressierten Eigenschaften in der Regel invariant im Hinblick auf Rotation und Skalierung der Bildinhalte oder auch auf Beleuchtungsveränderungen sein.

Häufig werden Bildeigenschaften nicht für das ganze Bild, sondern für Bildbereiche bestimmt, die z. B. mit Hilfe von Segmentierungsverfahren gewonnen wurden. Die Lokalität wird von neueren Verfahren wie SIFT (Scale Invariant Feature Transform) [13] noch weiter getrieben. Bilder werden hier durch eine Menge von Keypoints repräsentiert, die wiederum durch Histogramme zur Richtung der Gradienten im entsprechenden Bildpunkt repräsentiert werden. Damit ist z. B. eine Objektsuche recht gut möglich, weil man zu einem Bild des Anfrageobjekts nach Bildern in der Bildkollektion suchen kann, die möglichst viele Keypoints enthalten, die zu den Keypoints des Anfrageobjekts ähnlich sind.

Ein Ansatz bei der Bildsuche besteht darin, verschiedene Kriterien zu berücksichtigen. Bild 2 zeigt den Prototyp einer Bildsuche für mobile Geräte, der die Suche anhand von Farb- und Kontrasteigenschaften mit Se-



**Bild 2** Facettierte Bildsuche auf Basis von Bildinhalt, Datum, Tags und GPS-Koordinaten auf einem mobilen Endgerät [11].

lektionskriterien für Datum, Tags und Geo-Koordinaten verbindet. Der Nutzer kann in der Zeile am unteren Ende des Displays die gewünschten Selektionskriterien anwählen und bekommt dann im Stile einer facettierten Suche eine Auswahl von Facettenausprägungen (z. B. Farbbereiche oder Zeiten) angezeigt, die er zur Selektion auswählen kann.

## 4 Audiosuche

Die Suche in und nach Audiodateien arbeitet ähnlich wie die Bildsuche. Dabei unterscheidet man für die Repräsentation zwischen akustischen Features und semantischen Features. Diese müssen durch Transformation und Berechnung aus dem digitalisierten Audiosignal extrahiert werden. Zu den akustischen Features zählen die Lautstärke, die Energie in den Subbändern oder die Bandbreite. Semantische Features sind z. B. die Klangfarbe, der Rhythmus oder die genutzten Instrumente [20]. Einen Spezialfall bildet die Suche in MIDI-Daten (Musical Instrument Digital Interface), da hier die Repräsentation der Audiodaten auf einer semantisch höheren Ebene durch Tonhöhen und Notendauern erfolgt.

Anfragen werden im Audio-Retrieval häufig per QBE mit einem kurzen Audio-Fragment gestellt. In jüngerer Zeit sind erste Produkte auf dem Markt erschienen, mit

denen man sein mobiles Endgerät nutzen kann, um einen kurzen Mitschnitt eines gerade gehörten Audio-Titels an ein Musik-Portal zu senden und dort das entsprechende Musikstück zu erwerben. Eine andere Möglichkeit der Anfrageformulierung ist *Query by Humming*. Hier pfeift der Antragsteller das Suchmuster ins Mikrofon. Den Anfragen ist dabei gemein, dass die Beispielsequenz nicht zum kompletten Musikstück ähnlich ist, sondern nur zu einem kurzen Ausschnitt. Ein wichtiges Konzept der Audio-Repräsentation für die Suche ist daher das *Windowing*. Dabei wird die Audiodatei in kurze, überlappende Abschnitte unterteilt.

Weitere wichtige Fragestellungen des Audio-Retrieval sind die Segmentierung von Audiodaten, die Genre-Klassifikation, die Erkennung von Instrumenten oder die Umwandlung von gesprochener Sprache in Text [14; 20].

## 5 Videosuche

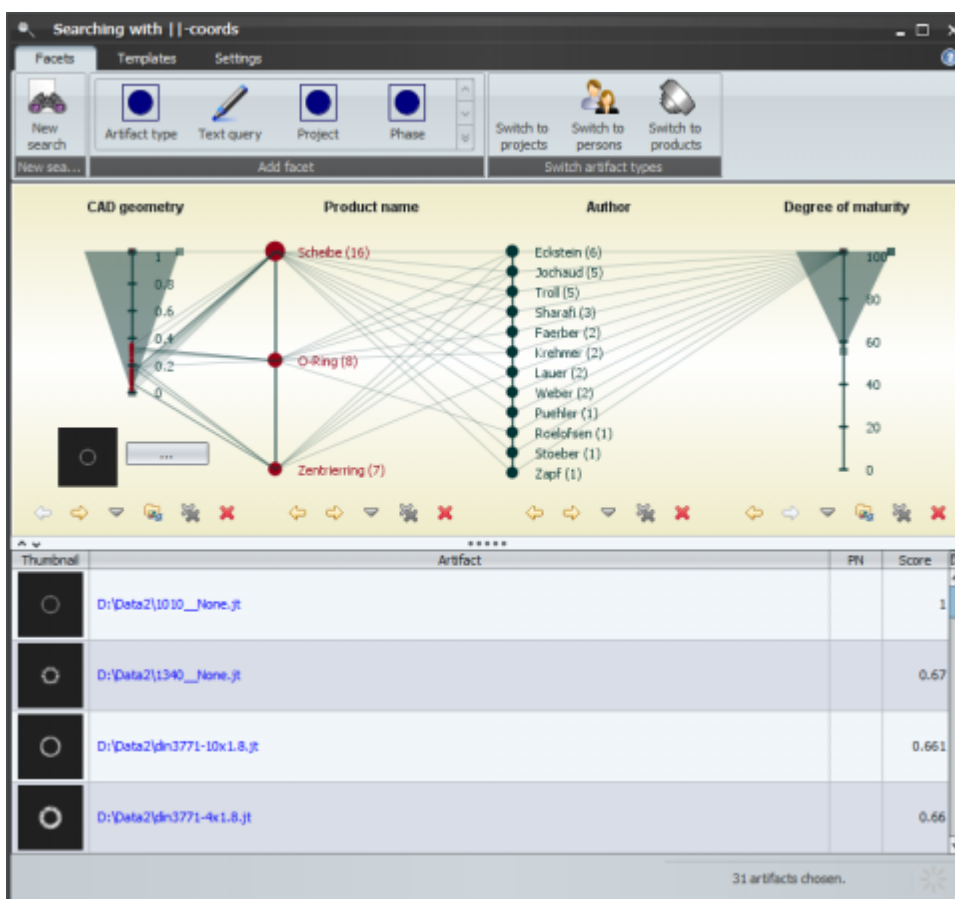
Die Suche in und nach Videodaten [21] kann als eine erste Stufe der Suche in multimedialen Daten gesehen werden, da ein Video aus einer Sequenz von Bildern, einem zugeordneten Audio sowie bisweilen auch Text in Form von Untertiteln besteht. Sofern der Text als solcher zugreifbar ist, bietet er eine gute Ausgangsbasis für die Suche. Auch die Audiodaten können mit den in Abschnitt 4 skizzierten Ansätzen adressiert werden.

Videodaten bieten aber auch zusätzliche Herausforderungen. Zunächst ist hier die Videostrukturanalyse zu nennen, die von der Shot-Erkennung über die Szenenerkennung bis zur Strukturierung der zugrundeliegenden Story reicht. Ein weiteres Feld bildet die Bestimmung von Schlüsselbildern, die bei der Suche adressiert werden können. Ebenso wichtig sind Schlüsselbilder für die Darstellung der Ergebnisse einer Videosuche.

Einen Eindruck von den Herausforderungen im Bereich der Videosuche können die Aufgabenstellungen geben, die derzeit bei der Evaluierungsinitiative TrecVID betrachtet werden. 2008 waren die Themenfelder: die Erkennung von Ereignissen in Überwachungsvideos, die Extraktion von High-Level-Beschreibungen wie „Klassenzimmer“, „Brücke“ ..., die Suche z. B. nach Personen, Dingen oder Orten, die Zusammenfassung von Film-Rohmaterial sowie die inhaltsbasierte Erkennung von verfremdeten Kopien [17].

## 6 Spezielle Objekttypen

Obwohl Texte, Bilder, Audioobjekte und Videos sicher die verbreitetsten Medientypen sind, ergeben sich gerade für speziellere Medientypen häufig interessante Suchszenarien. So spielt die Suche nach 3D-Objekten [3] in Anwendungsgebieten wie CAD, Virtual Reality, Medizin oder bei Computerspielen eine wichtige Rolle.



**Bild 3** Facettierte Suche in Kombination mit QbE auf 3D-Daten (CAD geometry), Selektionsbedingungen (Product name) und Nutzerpräferenzen (Degree of maturity) mit LFRP-Search [5].

Das Ziel ist in der Regel zu einem 3D-Anfrageobjekt von der Gestalt her ähnliche Objekte zu ermitteln (QbE). Zur Repräsentation der 3D-Objekte wurden verschiedene Verfahren vorgeschlagen. In einem einfachen Beispiel erzeugt man eine bestimmte Anzahl von gleichförmig verteilten Punkten innerhalb jedes Objekts und ermittelt die Abstände der Punkte untereinander. Über die Verteilung dieser Abstände bildet man ein Histogramm und geht davon aus, dass Objekte mit ähnlichen Histogrammen eine ähnliche Form haben [15].

Obwohl die auf der Form basierenden Ähnlichkeitsmodelle in speziellen Anwendungsfeldern gute Ergebnisse erzielen können, sieht man bei 3D-Objekten recht gut, dass oft eine Vielzahl von Faktoren für die Relevanz eines Objekts im Hinblick auf eine Anfrage von Bedeutung sind. Sucht man z. B. ein bestimmtes Bauteil, dann können Festigkeit, Gewicht, lieferbare Stückzahl und andere Faktoren neben der Form eine wichtige Rolle spielen. Man spricht in solchen Fällen von multikriteriellen Anfragen. Bild 3 zeigt den Prototyp einer Suchmaschine, die eine 3D-Ähnlichkeitssuche mit Selektionsbedingungen und Nutzerpräferenzen für Kriterien kombinieren kann. Die Linien im Suchfenster repräsentieren dabei die Datensätze im Ergebnis. Der Nutzer kann für einzelne Kriterien Selektionen vornehmen oder durch die grau hinterlegten Bereiche seine Präferenzen interaktiv definieren und so das Ranking im unteren Bereich des Tools beeinflussen.

## 7 Multimediale Daten

Während wir bisher davon ausgegangen sind, dass ein einzelnes Medienobjekt gesucht wird, ergeben sich durch

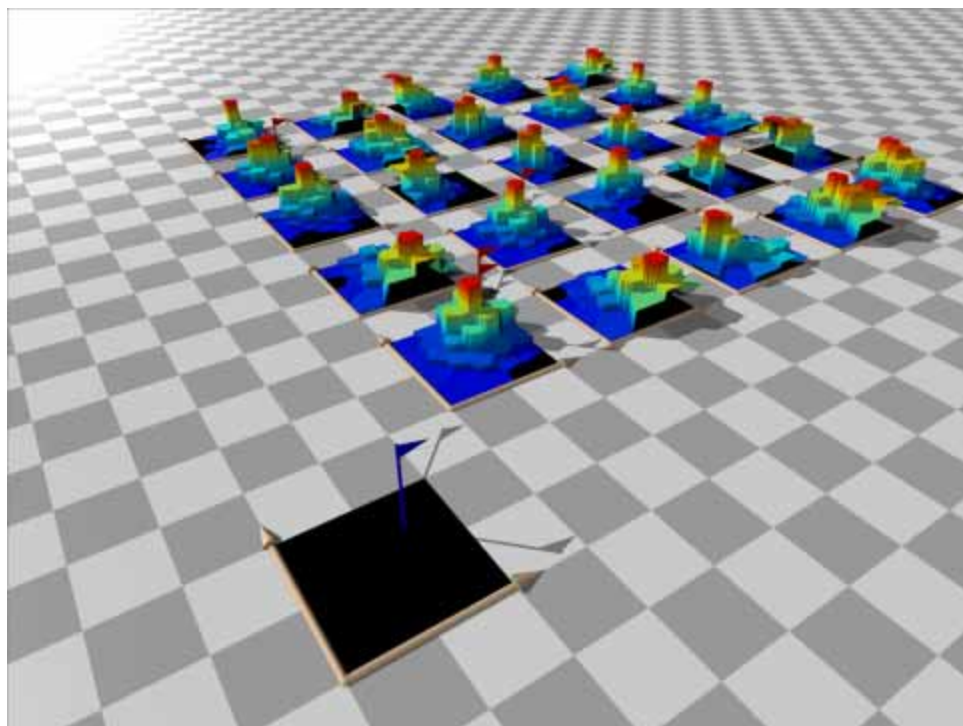
Formate wie XML, SMIL, MPEG-4, MPEG-7, ... neue Möglichkeiten und Herausforderungen. Die Verknüpfung von Medienobjekten zu Multimedia-Dokumenten erlaubt einerseits das Ausnutzen der Verbindungen. Nach einem Bild kann über den Text in der Umgebung gesucht werden. Auf der anderen Seite ergeben sich aber auch komplexe Fragen der Modellierung und der Anfrageformulierung [1]. Wie kann man beispielsweise eine Gewichtung mehrerer Anfragekriterien vornehmen? Sollten ganze Dokumente, einzelne Kapitel oder Medienobjekte in das Ergebnis aufgenommen werden? Dabei verschwimmen schnell die Grenzen zwischen *Suchen* und *Browsen*, weil das Ergebnis einer Anfrage der Ausgangspunkt zum weiteren Browsen in den gefundenen Dokumenten sein kann und andererseits ein durch Navigieren gefundenes Dokument Ausgangspunkt einer Anfrage („Gib mir alle Bilder, die von hier über Verknüpfungen eines bestimmten Typs erreicht werden können.“).

Ein Beispiel für ein System, das auf Basis von XML Strukturinformationen und Text zur Bildsuche nutzt, beschreibt [18].

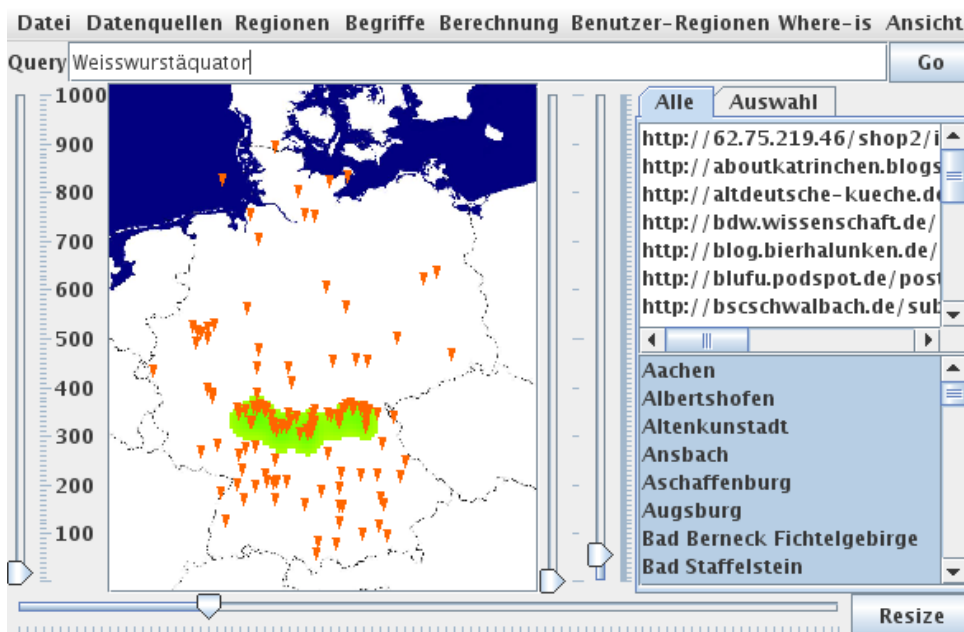
## 8 Aktuelle Trends

Unter den aktuellen Entwicklungstendenzen sollen an dieser Stelle nur vier herausgegriffen werden:

In Bild 1 ist bereits die Möglichkeit der Rückkopplung vorgesehen. Dies trägt der Tatsache Rechnung, dass die erste Formulierung eines Informationswunsches oft sehr grob ist. Durch eine *interaktive Suche* können deutliche Verbesserungen der Ergebnisqualität erzielt werden. Dabei muss die Interaktion komfortabel und zielgerichtet erfolgen, um beim Nutzer Akzeptanz zu finden. Ein Bei-



**Bild 4** Repräsentation der von einzelnen Peers in einem P2P-Netz verwalteten Daten durch Cluster-Histogramme [6].



**Bild 5** Bestimmung vager Regionen am Beispiel des Begriffs „Weißwurstäquator“ [10].

spiel für ein interaktives System zeigt Bild 3. Hier wird dem Benutzer im Rahmen einer facettierten Suche verdeutlicht, wie groß die verbleibende Ergebnismenge bei der Wahl weiterer Selektionskriterien wäre.

Ein zweiter Bereich ist die *Suche in verteilten Systemen*, bei der mehrere Quellen betrachtet werden. Hier ergeben sich offene Fragen bei der Wahl der Quellen (*Source Selection*) und bei der Integration der Ergebnisse mehrerer Quellen zu einem Gesamtergebnis (*Fusion*). Bild 4 verdeutlicht die kompakte Beschreibung der Daten der Peers in einem Peer-to-Peer-Netzwerk mit Hilfe von Clusterhistogrammen (dabei werden zweidimensionale Repräsentationen der Medienobjekte angenommen). Für eine gegebene Anfrage kann auf Basis dieser Histogramme entschieden werden, welche Peers in die Anfragebearbeitung einbezogen werden sollten.

Einen dritten Bereich bildet die *kollaborative Suche*. Man unterscheidet dabei Situationen, in denen eine Gruppe gezielt gemeinsam an einer Recherche arbeitet (synchrone Kollaboration) und Situationen, in denen z. B. die Mitarbeiter eines Unternehmens von den Vorarbeiten anderer profitieren können („Andere Nutzer in einer ähnlichen Situation fanden folgende Treffer hilfreich...“).

Schließlich verfügen mehr und mehr Medienobjekte über geografische Metadaten. Dies erlaubt eine präzise Suche z. B. nach Bildern vom Kölner Dom. Andererseits bereiten umgangssprachliche Formulierungen bisweilen Probleme. Bild 5 zeigt ein Werkzeug, mit dem es auf Basis von Textanalysen und statistischen Auswertungen möglich ist, die Lage umgangssprachlicher Bezeichnungen näherungsweise zu bestimmen.

Die Zukunft des Multimedia Information Retrieval liegt dabei wohl unter anderem in Suchsystemen, denen es gelingt, durch Einbeziehung von Kontext- und

Domänenwissen die semantische Lücke ansatzweise zu überwinden.

## Literatur

- [1] Anastasia Analyti and Stavros Christodoulakis. *Content-based querying*. In: Peter M. G. Apers, Henk M. Blanken, and Maurice A. W. Houtsma, editors, *Multimedia Databases in Perspective*, pages 145–179. Springer, 1997.
- [2] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. *Modern Information Retrieval*. Addison Wesley, May 1999.
- [3] Benjamin Bustos, Daniel A. Keim, Dietmar Saupe, Tobias Schreck und Andrada Tatu. *Methoden und Benutzerschnittstellen für effektives Retrieval in 3D-Datenbanken*. In: *Datenbank-Spektrum*, 7(20):23–32, 2007.
- [4] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Ze Wang. *Image retrieval: Ideas, influences, and trends of the new age*. In: *ACM Comput. Surv.*, 40(2), Article No. 5, 2008.
- [5] Raiko Eckstein and Andreas Henrich. *Visual browsing in product development processes*. In: *Proc. 17th Int'l Conf. on Engineering Design*, Stanford, CA, 2009. The Design Society.
- [6] Martin Eisenhardt, Wolfgang Müller, Andreas Henrich, Daniel Blank, and Soufyane El Allali. *Clustering-based source selection for efficient image retrieval in peer-to-peer networks*. In: *ism*, pp. 823–830, Eighth IEEE Int'l Symposium on Multimedia (ISM '06), 2006.
- [7] David Feng, W. C. Siu, and Hong J. Zhang, editors. *Multimedia Information Retrieval and Management – Technological Fundamentals and Applications*. Signals and Communication Technology. Springer, Berlin, 2003.
- [8] Reginald Ferber. *Information Retrieval – Suchmodelle und Data-Mining-Verfahren für Textsammlungen und das Web*. dpunkt-verlag, Heidelberg, 2003.
- [9] Alaa H. Halawani, Alexandra Teynor, Lokesh Setia, Gerd Brunner, and Hans Burkhardt. *Fundamentals and Applications of Image Retrieval: An Overview*. In: *Datenbank-Spektrum*, 6(18):14–23, 2006.
- [10] Andreas Henrich, Volker Lüdecke, and Daniel Blank. *Approaches for determining the geographic footprint of arbitrary terms for retrieval and visualization*. In: *GIS '08: Proc. of the 16th ACM SIGSPATIAL Int'l Conf. on Advances in Geographic Information*

- Systems, pp. 1–4, New York, NY, USA, 2008. ACM.
- [11] Adrian Hub, Daniel Blank, Andreas Henrich, and Wolfgang Müller. *Picadomo: Faceted image browsing for mobile devices*. In: Proc. of the Seventh Int'l Workshop on Content-Based Multimedia Indexing, pp. 249–254, Chania, Crete Island, Greece, June 2009. IEEE Computer Society.
- [12] Bernd Jähne. *Digitale Bildverarbeitung*. Springer, Berlin Heidelberg, 6. überarbeitete und erweiterte Auflage, 2005.
- [13] David G. Lowe. *Distinctive image features from scale-invariant keypoints*. In: Int'l Journal of Computer Vision, 60(2):91–110, 2004.
- [14] Meinard Müller, Frank Kurth und Michael Clausen. *Aktuelle Aspekte des Music Information Retrieval*. In: Datenbank-Spektrum, 6(18):24–32, 2006.
- [15] Robert Osada, Thomas A. Funkhouser, Bernard Chazelle, and David P. Dobkin. *Shape distributions*. In: ACM Trans. Graph., 21(4):807–832, 2002.
- [16] Stephen Robertson, Hugo Zaragoza, and Michael Taylor. *Simple BM25 extension to multiple weighted fields*. In: CIKM '04: Proc. of the 13th ACM Int'l Conf. on Information and Knowledge Management, pp. 42–49, New York, NY, USA, 2004. ACM.
- [17] Alan F. Smeaton, Paul Over, and Wessel Kraaij. *Evaluation campaigns and trecvid*. In: MIR '06: Proc. of the 8th ACM Int'l Workshop on Multimedia Information Retrieval, pp. 321–330, New York, NY, USA, 2006. ACM Press.
- [18] Mouna Torjmen, Karen Pinel-Sauvagnat, and Mohand Boughanem. *XML multimedia retrieval: From relevant textual information to relevant multimedia fragments*. In: Advances in Information Retrieval, Proc. of the 31th European Conf. on IR Research, ECIR 2009, Toulouse, France, April 6–9, 2009, LNCS 5478, pp. 150–161. Springer, 2009.
- [19] Julia Vogel, Adrian Schwaninger, Christian Wallraven, and Heinrich H. Bühlhoff. *Categorization of natural scenes: Local versus global information and the role of color*. In: ACM Trans. Appl. Percept., 4(3):19, 2007.
- [20] Changsheng Xu, David Dagan Feng, and Qi Tian. *Content-Based Retrieval for Digital Audio and Music*. In: Feng et al. [7], Chapter 5, pp. 95–120, 2003.
- [21] Hongjiang Zhang. *Content-Based Video Analysis, Retrieval and Browsing*. In: Feng et al. [7], Chapter 2, pp. 27–56, 2003.

Manuskripteingang: 19. September 2009



**Prof. Dr. Andreas Henrich** erhielt sein Diplom in Wirtschaftsinformatik 1987 von der TU Darmstadt, den Grad eines Dr. rer. nat. 1990 von der FernUniversität Hagen und die Venia Legendi in Informatik 1997 von der Universität Siegen. 1998 wurde er auf eine Professur für Praktische Informatik an der Otto-Friedrich-Universität Bamberg berufen. Von 2001 bis 2004 war er Inhaber des Lehrstuhls Angewandte Informatik I an der Universität Bayreuth und seit 2004 ist er Inhaber des Lehrstuhls Medieninformatik in der Fakultät Wirtschaftsinformatik und Angewandte Informatik der Otto-Friedrich-Universität Bamberg. Seine Forschungsschwerpunkte liegen in den Bereichen: Information Retrieval, Datenvisualisierung und -exploration, multimediale Systeme sowie E-Learning.

Adresse: Otto-Friedrich-Universität Bamberg, Feldkirchenstraße 21, 96052 Bamberg, Deutschland, Tel.: +49-951-863-2850, Fax: +49-951-863-2852, E-Mail: andreas.henrich@uni-bamberg.de