

PythiaSearch - Interaktives, Multimodales Multimedia-Retrieval

David Zellhöfer, Thomas Böttcher, Maria Bertram, Christoph Schmidt,
Claudius Tillmann, Markus Uhlig, Marcel Zierenberg, Ingo Schmitt

Brandenburgische Technische Universität
Walther-Pauer-Str. 1, 03046 Cottbus
david.zellhoefer|tboettcher@tu-cottbus.de

Abstract: *PythiaSearch* ist ein interaktives Multimedia-Retrieval-System. Es vereint verschiedene Suchstrategien, diverse Visualisierungen und erlaubt eine Personalisierung der Retrieval-Ergebnisse mittels eines Präferenz-basierten Relevance Feedbacks. Das System nutzt die probabilistische Anfragesprache *CQQL* und erlaubt eine multimodale Anfragedefinition basierend auf Bildern, Texten oder Metadaten.

1 Motivation

Multimodale Retrievalsysteme (MIRS) sind häufig nur beschränkt anpassbar. Meistens können nur Gewichtungen von fest vorgegebenen Features verschoben werden, um die Anfrage an das Informationsbedürfnis des Nutzers anzupassen. *PythiaSearch* stellt einen adaptiveren Ansatz dar, welcher auf der probabilistischen, logikbasierten Anfragesprache CQQL [Sch08] basiert. Die Stärke des Systems liegt dabei vor allem in der Kombinationen von ähnlichkeitsbasierten und booleschen Anfragebedingungen, deren Gewichtung mittels Präferenzen angepasst werden kann. Hierdurch wird insbesondere Experten ein personalisierbares Werkzeug geboten, welches auf die volle Mächtigkeit einer logikbasierten Anfragesprache zurückgreifen kann.

2 Schnittstelle und Interaktion

Die graphischen Schnittstelle (GUI) unterstützt verschiedene Suchstrategien (gerichtet und explorativ), die während typischen Suchprozessen zu beobachten sind [RMMH00]. Dabei wird durchgängig auf die Anfragesprache CQQL zurückgegriffen, welche das kognitive Retrievalmodell der Polyrepräsentation [Ing96] umsetzt. Hierdurch wird es möglich, die GUI und die Anfrageverarbeitung ohne konzeptionelle Brüche umzusetzen [Zel12b]. In dieser Arbeit soll die graphische Oberfläche vorgestellt und unterschiedlichen Suchstrategien, Ergebnisvisualisierungen sowie dazugehörige Personalisierungsmöglichkeiten in Erweiterung von [ZBB⁺12] erläutert werden. Gemäß der Prinzipien der nutzerzentrierten Softwareentwicklung wurden die Anforderungen an die Software in Kooperation mit po-

tentiellen Nutzer der Medien- und Marktforschungsbranche (z.B. Bertelsmann, Deutsche Telekom oder TNS Infratest) im Rahmen von drei Workshops in 2011 und 2012 erhoben. Die Leistungsfähigkeit konnte in einer Nutzungsstudien [Zel12a] gezeigt werden.

Aufbau der grafischen Benutzeroberfläche Der vorgestellte Prototyp ist für die gängigen Betriebssysteme Mac OS X, Windows sowie Linux verfügbar und ermöglicht dem Nutzer die direkte Interaktion mit den visualisierten Dokumenten (z.B. Bildern, PDFs, etc.). Im Folgenden sollen zunächst die Grundelemente der GUI (siehe Abbildung 1) beschrieben werden.



Abbildung 1: Aufbau der GUI (Mac OS X)

1. Das Eingabe-Feld dient zur Vergabe von Suchwörtern. Diese können z.B. mittels boolescher Operatoren verbunden werden.
2. Das multimediale Eingabefenster ermöglicht es dem Anwender ein oder mehrere QBE-Dokumente (Query By Example; z.B. ein Bild oder PDF) zu wählen.
3. Die Steuerung für die Suche kann genutzt werden, um eine neue Suche zu starten sowie die Anzahl der angezeigten Dokumente zu konfigurieren.
4. Der Suchverlauf erlaubt es dem Anwender, bereits durchgeführte Suchen, Lernschritte etc. wieder aufzurufen bzw. wieder rückgängig zu machen.
5. In der Ergebnissicht werden alle relevanten Dokumente dargestellt. Diese Sicht erlaubt dem Nutzer eine direkte Interaktion mit den visualisierten Dokumenten. Hierbei können diese verschoben als auch gestapelt um z. B. für die entstandene Gruppe Annotationen zu vergeben.
6. Mithilfe dieses Auswahlmenüs können verschiedene Visualisierungen der Ergebnissicht (5) eingestellt werden. In Abbildung 1 (zentrales Fenster) ist die Matrix-Ansicht dargestellt, welche die Elemente nach absteigender Relevanz sortiert. Weitere Details zu den Visualisierungen finden sich in Abschnitt 2.
7. Für eine Personalisierung der Suchergebnisse können Dokumente aus der Visualisierung in das Präferenz-Fenster (mithilfe von Drag & Drop oder dem Kontextmenü) gezogen werden. Hierbei wird eine Halbordnung definiert, welche die Relevanz der Objekte in Bezug auf die Anfrage beschreibt. Das QBE-Dokument befindet sich hierbei im Zentrum des Fensters und die Relevanz der Dokumente nimmt mit steigender Entfernung zum Zentrum ab [Zel12b].

8. Die facettrierte Suche erlaubt es, einen Filter auf die bisherigen Suchergebnisse zu setzen. Hierbei spiegelt eine Facette eine boolesche Bedingung wieder, welche direkt in eine gewichtete CQQL-Anfrage transformiert wird. In der Ergebnisliste werden bei Aktivierung einer Facette nur Dokumente gezeigt, welche die definierte Bedingung erfüllen (z. B. die Abwesenheit von Personen auf einem Foto).
9. In dieser Ansicht werden repräsentative Bilder der durchsuchten Datenbank angezeigt, um diese explorativ erschließen zu können. Bei Auswahl eines der Bilder werden ähnliche Bilder in der jeweils ausgewählten Visualisierung angezeigt, um Browsing zu ermöglichen.

Unterstützte Suchstrategien *PythiaSearch* unterstützt gängige Suchstrategien. Hierbei wird eine *gerichtete Suche*, bei der der Nutzer bereits sein Informationswunsch kennt, als auch eine *explorative Suche* ermöglicht. Beide Suchstrategien sind kombinierbar, wobei der Wechsel der Suchstrategie jederzeit aus einer beliebigen Ansicht vorgenommen werden kann. Für die *gerichtete Suche* stehen dem Nutzer zwei Eingabefelder zur Verfügung. Mit dem Texteingabefeld können einfache Keyword-basierte Suchanfragen definiert werden. Über das multimediale Eingabefenster können sowohl Bilder als auch PDF-Dokumente zur Anfragedefinition genutzt werden. Beide Eingabe können kombiniert werden, so dass ein Informationswunsch auf multimodaler Ebene definiert werden kann. Enthält ein PDF-Dokument neben Text auch Bilder so wird die gesamte Struktur zur Ähnlichkeitsberechnung herangezogen (sowohl Bilder als auch Text in Abhängigkeit der Struktur des Dokuments). Für jede vorhandene Dokumentenrepräsentation auf jeder Ebene eines jeden Anfragedokuments wird eine Ähnlichkeitsberechnung durchgeführt. Über einen speziellen Operator (z.B. auf CQQL basierend) werden die Einzelähnlichkeiten miteinander aggregiert. Für Anwender, die ihren Informationswunsch nicht explizit definieren können wurde das *explorative Browsing* integriert. Es ermöglicht einen Überblick über einzelne Dokumente innerhalb der verwendeten Datenbank zu erhalten (siehe Abbildung 1 (9)). Die Datenbank wird dabei mit einem Cluster-basierten Ansatz aufgearbeitet, so dass dem Anwender zunächst nur ein Element einer Klasse präsentiert wird. Über die einzelnen Visualisierungsformen können dann die Objekte innerhalb einer Klasse visualisiert werden. Die während der Exploration gefundenen Dokumente können im Anschluss beispielsweise als QBE-Dokument genutzt werden, um eine gerichtete Suche zu starten.

Ergebnisvisualisierung

Ausgehend von den verschiedenen Suchstrategien können aktuell drei unterschiedliche Visualisierungen gewählt werden, um die Ergebnisdokumente zu präsentieren.

Matrix In der Standardansicht werden Dokumente durch die berechneten Ähnlichkeitswerte sortiert und in einem Raster angezeigt. Diese Ansicht ist ideal, um das Ranking einer Suchanfrage zu betrachten und wird durch eine gerichtete Suchanfrage generiert.

SOM Die SOM-Ansicht ist eine selbst-organisierende Karte [Koh95] in der alle Objekte durch die gewählte Eigenschaft (wie z. B. Farbe, Textur oder einer CQQL-Anfrage) automatisch sortiert werden. Die SOM ermöglicht es, einen Überblick über die verschiedenen

Ausprägungen aller Medienelemente zu erhalten, hierbei ist es möglich sich auf eine Ausprägung zu fixieren (z.B. eine konkrete Farbe) und in diesen Bereich hinein zu navigieren. Dokumente, die sich innerhalb der SOM nahe beieinander liegen besitzen eine ähnliche Charakteristik in der gewählten Eigenschaft.

Cluster Die Ergebnisse einer Suche können durch das Clustern weiter verarbeitet werden, um neue Eigenschaften aufzuzeigen oder ähnliche Elemente zu gruppieren. Im Gegensatz zur SOM erfolgt die Trennung hier hart, d.h. Die Ergebnisse gehören zu genau einem Cluster. Ein Cluster beinhaltet Dokumente, die bezüglich einer CQQL-Formel möglichst ähnlich zueinander sind, während Elemente unterschiedlicher Cluster möglichst unähnlich von einander sind. Die gewählte CQQL-Formel kann ein Merkmal aber auch eine beliebige (logische) Kombination von Merkmalen (z. B. GPS-Koordinaten und Farbe) verwenden, wodurch die definierte Eigenschaft für das Clustern sehr flexibel ist und manuell auf die Bedürfnisse des Nutzers angepasst werden kann.

3 Demonstration

Innerhalb der Demonstration soll der typische Ablauf eines Retrieval-Prozesses gezeigt werden. Hierbei wird die Extraktion von Features, die Auswahl eines Anfragedokumentes und die verschiedenen Visualisierungen (vgl. Abschnitt 2) vorgestellt. Desweiteren wird die Verfeinerung eines Informationswunsches durch die Vergabe von Präferenzen durchgeführt und somit eine Personalisierung der Ergebnisse erzielt. Im weiteren Verlauf wird durch eine kombinierte Anfrageformulierung eine multimodale Suche demonstriert.

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 03FO3072 gefördert.

Literatur

- [Ing96] Peter Ingwersen. Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory. *Journal of Documentation*, 52:3–50, 1996.
- [Koh95] Teuvo Kohonen. *Self-organizing maps*. Jgg. 30 of *Springer series in information sciences*. Springer, Berlin, 1995.
- [RMMH00] Harald Reiterer, Gabriela Mußler, M. Thomas Mann und Siegfried Handschuh. INSYDER - an information assistant for business intelligence. In *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '00, Seiten 112–119. ACM, 2000.
- [Sch08] Ingo Schmitt. QQL: A DB&IR Query Language. *The VLDB Journal*, 17(1):39–56, 2008.
- [ZBB⁺12] David Zellhöfer, Maria Bertram, Thomas Böttcher, Christoph Schmidt, Claudius Tillmann und Ingo Schmitt. PythiaSearch – A Multiple Search Strategy-supportive Multimedia Retrieval System. In *Proceedings of the 2nd ACM International Conference on Multimedia Retrieval*, ICMR '12, Seite to appear. ACM, 2012.
- [Zel12a] David Zellhöfer. On the Usability of PythiaSearch. Bericht 9, Brandenburg University of Technology, Cottbus, 2012.
- [Zel12b] David Zellhöfer. A permeable expert search strategy approach to multimodal retrieval. In *Proceedings of the 4th Information Interaction in Context Symposium*, IIIIX '12, Seiten 62–71, New York, NY, USA, 2012. ACM.