Interoperable and Unified Multimedia Retrieval in Distributed and Heterogeneous Environments

Florian Stegmaier  
Chair of Distributed Information Systems, University of Passau, Germany  
florian.stegmaier@uni-passau.de  
Supervisor: Harald Kosch, Co-Supervisor: Mario Döller

ABSTRACT
In this abstract, the research topics of my doctoral thesis will be introduced. These emerged within THESEUS, in which I work as a third-party funded researcher. The overall aim of my work is to provide unified and interoperable multimedia retrieval in distributed and highly heterogeneous environments. In this case, I am focusing on multimedia databases, multimedia metadata formats, semantic web technologies as well as international standardization (MPEG/JPEG and W3C).

Categories and Subject Descriptors
H.2.5 [DATABASE MANAGEMENT]: Heterogeneous Databases; H.5.1 [INFORMATION INTERFACES AND PRESENTATION]: Multimedia Information Systems

General Terms
Experimentation, Standardization, Design

Keywords
Multimedia retrieval, heterogeneous databases, interoperability, multimedia metadata

1. INTRODUCTION
Nowadays, the public interest in multimedia retrieval is tremendous. This trend is based on several facts: First, anybody has the ability to produce digital multimedia content in an easy fashioned way. Almost every mobile device is already equipped with image sensors for taking still images or even movies. Along with this, the acquisition costs, like for digital cameras, decrease dramatically. Second, there are uncountable available services for publishing this content: Flickr for still images, Youtube for movies, Slideshare for presentation materials, and so on. Finally, the gap between production and consumption of multimedia data is closed by cheap and nearly everywhere accessible high-speed internet infrastructures (e.g., UMTS).

Multimedia retrieval still features unsolved issues [4], such as the Semantic Gap or interoperability issues [7]. Since interoperability has many faces, my research activities are dealing with the interoperability issues among different query languages, metadata formats and the actual data representation / serialization.

2. CENTRAL RESEARCH TOPICS
This section introduces my research questions. Examples about the major issues, related work, possible solutions along with the actual state of my research will be sketched in Section 3.

(i) How to provide an unified access to distributed and heterogeneous environments?
(ii) How can query semantic be maintained during metadata and query language transformation process?
(iii) Can query classification and backend benchmarking establish an efficient / intelligent data fusion process?
(iv) How can heterogeneous multimodal multimedia data be accessed and visualized?

3. STATE OF RESEARCH
The initial stage of research concentrated on the unified access issue (i), as it builds the basis for the remaining three research tasks. Hereby, crucial tasks were the choice of an appropriate query language and the identification of key-components inside the specified framework. The outcome of this investigation is the AIR multimedia middleware framework, presented in [9]. It makes use of the recently issued multimedia specific query language MPEG Query Format (MPQF), because of its adjustment to general multimedia retrieval needs, as shown in [3]. More specific, in the terms of image retrieval AIR implements the JPEG Query Format (subset of MPQF), which is defined in the JPSearch project. AIR is able to support both, distributed query processing, as well as local query processing. Based on this, the architecture of AIR is divided in several (logically independent) components and contains placeholders, that have not yet been implemented. The development of these have a strong corellation with the remaining research questions. The highlights of the current implementation are service discovery functionalities for automatic query segmentation, providing an unified access to distributed databases using MPQF as an abstraction layer and metadata transformation functionalities. At the moment, a preliminary version of AIR has been already integrated in scientific projects (e.g., THESEUS: Medico or an interoperable image search engine).

The metadata transformation process (ii) implements the JPSearch transformation rules [2]. They define XML-based syntactical mappings between the JPSearch core schema and an arbitrary (XML based) metadata format. The present
implementation is able to transform metadata informations inside an isolated, metadata specific query type. The current approach will be enlarged in order to keep the original query semantic while changing the complete query structure to avoid transformation conflicts, as shown in Figure 1. In addition, I am also involved in the W3C Media Annotations Working Group. Here, an ontology based transformation process \cite{8} is proposed in order to define syntactic and semantic mappings among metadata formats.

As the given references indicate, research topics (i) and (ii) are currently under investigation and first results are already available. On the other hand, research topics (iii) and (iv) are in a very early stage.

Recent research efforts dealt with the collection fusion problem, like \cite{1}, \cite{10} or \cite{5}. Yet available algorithms used in an aggregation process are mostly tied to the type of the underlying data (e.g., multimedia data type or metadata format). I want to establish an intelligent result aggregation (iii), which is able to select the best set of algorithms of an algorithm pool. The input parameters for this process will be the output of a MPQF/JPQF query classification (e.g., specifying the targeted features) and a backend benchmarking process. This process collects important informations about the retrieval quality of a backend, since every database uses its own retrieval algorithms, low–level features (such as MPEG-7 descriptors) and quality measures. Here, a generic and parameterizable evaluation function should be investigated, which is not bound to a specific data set. The result of this evaluation function will serve as a measure for the actual retrieval quality of a backend and will be used for example in the reranking process. Finally, it has to be verified, whether relevance feedback, semantic web\footnote{http://www.w3.org/2001/sw/} or linked data\footnote{http://www.linkeddata.org/} techniques could provide a valuable input for these tasks.

The accessibility of multimodal data (iv) depends on two major aspects: efficient algorithms or data structures (e.g., indices) can heavily improve the quality of search engines, see for example VIDI–Video\footnote{http://www.vidivideo.info/}. In this project, several informations from different sources will be combined in order to create a semantic video search engine. In contrast to that, attractive and easy to operate user interfaces (e.g., FreeEye \cite{6}) may also improve the quality or the interaction abilities of an user by telling the system what exactly he is searching for. As part of the interoperable image search, we implemented an user interface providing an easy way of generating semantically rich queries. Cutting edge technologies like JavaFX are used for building these environments. Here, we are trying to adapt the yet available research solutions to the overall interoperability issues we are facing.

\section{Conclusion & Outlook}

This abstract gives an brief insight into the research challenges, I am facing in my PhD. The main goal of this thesis is the improvement of the interoperability issues that arise in distributed, heterogeneous and multimodal environments.

We started by considering two real world use cases (THE-SEUS:Medico and interoperable image search) serving as an input for our research. These achievements of our work can be inferred from the aforementioned integration in several projects and thus serve as a proof of concept.

Next steps to be done are further investigation of the data fusion issue as well as an evaluation, how user interfaces could be designed / implemented to present multimodal data in a user–friendly way.

\section{Acknowledgments}

I would like to thank my wife, Tamara, my supervisor, Harald Kosch, my co-supervisor, Mario Döller, my colleagues at our chair and all others, who always give me advice and support whenever needed.

\section{References}

\begin{thebibliography}{9}
\end{thebibliography}