The MPEG-7 query format: a new standard in progress for multimedia query by content

Kevin Adistambha  
*University of Wollongong, ka07@uowmail.edu.au*

M. Doeller  
*University of Passau, Germany*

R. Tous  
*Universitat Politècnica de Catalunya, Spain*

M. Gruhne  
*Fraunhofer Institute for Digital Media Technology IDMT, Germany*

M. Sano  
*NHK Science & Technical Research Laboratories, Japan*

*See next page for additional authors*

Publication Details

The MPEG-7 query format: a new standard in progress for multimedia query by content

Abstract
In recent years, the amount of Internet accessible digital audiovisual media files has vastly increased. Therefore the need to describe the media (by way of metadata) has also increased significantly. MPEG-7 (finalized in 2001) provides a comprehensive and rich metadata standard for the description of multimedia content. Unfortunately, a standardized query format does not exist for MPEG-7, or other multimedia metadata. Such a standard would provide for communications between querying clients and databases, supporting cross-modal and cross-media retrieval. The ISO/IEC SC29WG11 committee decided therefore to contribute to this application space by adding such functionality as a new part of the MPEG-7 series of standards. In response to a Call for Proposals, six proposals were submitted. This paper describes the strengths of each proposal as well as the resulting draft standard for the MPEG-7 query format.

Disciplines
Physical Sciences and Mathematics

Publication Details

Authors

This conference paper is available at Research Online: http://ro.uow.edu.au/infopapers/614
The MPEG-7 Query Format: A New Standard in Progress for Multimedia Query by Content

Kevin Adistambha, Mario Doeller, Rubén Tous, Matthias Gruhne, Masanori Sano, Chris Tsinaraki, Stavros Christodoulakis, Kyoungro Yoon, Christian H. Ritz and Ian S. Burnett

1School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Australia
E-mail: {ka07, critz, ianb}@uow.edu.au
2Dept. Distributed Information Technology, University of Passau, Innstrasse 43, Passau, Germany
E-mail: Mario.doeller@uni-passau.de
3Universitat Politècnica de Catalunya (UPC), Dpt. d’Arquitectura de Computadors, Barcelona, Spain
E-mail: rtous@ac.upc.edu
4Department Metadata, Fraunhofer IDMT, Germany
E-mail: ghe@idmt.fraunhofer.de
5NHK Science & Technical Research Laboratories, Japan
E-mail: sano.m-f@nhk.or.jp
6Lab. of Distributed Multimedia Information Systems & Applications (MUSIC), Technical University of Crete (TUC)
E-mail: {chrisa, stavros}@ced.tuc.gr
7School of Computer Science and Engineering, Konkuk University, 1 Hwayang-Dong, Kwangjin-Gu, Seoul, Korea
Email: yoonk@konkuk.ac.kr

Abstract—In recent years, the amount of Internet accessible digital audiovisual media files has vastly increased. Therefore the need to describe the media (by way of metadata) has also increased significantly. MPEG-7 (finalized in 2001) provides a comprehensive and rich metadata standard for the description of multimedia content. Unfortunately, a standardized query format does not exist for MPEG-7, or other multimedia metadata. Such a standard would provide for communications between querying clients and databases, supporting cross-modal and cross-media retrieval. The ISO/IEC SC29WG11 committee decided therefore to contribute to this application space by adding such functionality as a new part of the MPEG-7 series of standards. In response to a Call for Proposals, six proposals were submitted. This paper describes the strengths of each proposal as well as the resulting draft standard for the MPEG-7 query format.

The culmination of the research performed in the area of multimedia description by content is the finalization of the first MPEG-7 standard in 2001. This standard defines a way to describe the content of a multimedia content (e.g. audio, video and images) using both subjective and objective descriptors (metadata) using XML. However, while many useful descriptors have been standardized, what is lacking is a standardized way of searching for multimedia content. The ISO/IEC SC29WG11 committee decided therefore to contribute to this application space by adding such functionality as a new part of the MPEG-7 series of standards. In response to a Call for Proposals, six proposals were submitted. This paper describes the strengths of each proposal as well as the resulting draft standard for the MPEG-7 query format.

I. INTRODUCTION

The culmination of the research performed in the area of multimedia description by content is the finalization of the first MPEG-7 standard in 2001. This standard defines a way to describe the content of a multimedia content (e.g. audio, video and images) using XML. However, while many useful descriptors have been standardized, what is lacking is a standardized way of searching for multimedia content. The ISO/IEC SC29WG11 issued a call for proposal for the MPEG-7 Query Format (M7QF). Several proposals were submitted with the best features of each subsequently combined into the current M7QF draft standard.

This paper presents the status of the ongoing work in MPEG toward the final M7QF, and is organized as follows. Section II provides the overview of currently available technologies in the area of multimedia querying. Section III will briefly describe the call for proposal and the motivation for the development of the new standard, Section IV will describe proposals submitted to MPEG, Section V will provide a general overview of the current M7QF, and Section VI will provide the conclusions.

II. RELATED WORK TO MPEG-7 QUERY LANGUAGES AND DATABASES

A very good overview of the usability of XML databases for MPEG-7 is provided by Westermann and Klas in [1]. To summarize their findings, neither native XML databases (e.g., Tamino [2], Xindice [3]) nor XML database extensions (e.g., Oracle XML DB [4], Monet XML [5], etc.) provide full support for managing MPEG-7 descriptions with respect to their given requirements. In addition, currently available multimedia query languages in this domain show various limitations, either in respect of multimedia query operations or in supporting requests to different distributed database systems.

In the following, we will have a closer look to currently available query languages in various approaches and evaluate them in respect to its use for MPEG-7. As MPEG-7 bases on XML schema, one dents to use XML based query languages such as e.g., XPath [6], XQuery [7] or SQL/XML [8]. In general, XPath provides access to individual parts of data elements in XML documents but supports no means for e.g., grouping or joins. In addition, XPath on its own provides no functionality for querying multimedia data in MPEG-7 descriptions by any multimedia query operation. XQuery is a declarative query language, integrates XPath 2.0 and provides a rich set of operations and tools for querying XML descriptions. Besides, XQuery does not provide means for querying multiple databases in one request and does not
support multimodal or spatial/temporal queries. Nevertheless, there is ongoing work in this direction. For instance, the authors in [18] describe a XQuery extension for MPEG-7 vector-based feature queries. Furthermore, the authors in [19] adapted XQuery for the retrieval of MPEG-7 descriptions based on semantic views. Its adaptation, called Semantic Views Query Language (SVQL) is specialized for retrieving MPEG-7 descriptions in TV news retrieval applications and is not intended to be a general query language.

The authors in [9] propose an XML query language with multimedia query constructs called MMDOC-QL. MMDOC-QL bases on a logical formalism path predicate calculus [10] which supports multimedia content retrieval based on described spatial, temporal and visual data types and relationships. Of all investigated query languages, MMDOC-QL fulfills best retrieval in MPEG-7 descriptions. Nevertheless, there are several drawbacks such as simultaneous searches in multiple databases or the integration of user preferences and usage history which are not considered in MMDOC-QL.

XIRQL [11] is a query language for information retrieval in XML documents and bases on XQL [12]. The query language integrates new features that are missing in XQL such as weighting and ranking, relevance-oriented search, data types and vague predicates and semantic relativism. A similar weighting and relevance approach has been introduced in [17]. Besides, there exist several query languages explicitly for multimedia data such as SQL/MM [14], MOQL [15], POQLMM [16] etc. which are out of scope of this paper based on its limitation in handling XML data. In addition, there are some combined solutions (e.g., PTDOM [13], MPEG7-MMDB [20]) that provide retrieval and storage mechanism in one system.

III. REQUIREMENTS FOR THE MPEG-7 QUERY FORMAT

The first call for proposal for M7QF was issued in the 76th MPEG meeting in Montreux, Switzerland in April 2006 [21], with the final call for proposal issued in the 78th MPEG meeting in Hangzhou, China in October 2006. Fig. 1 illustrates the normative part of the proposed standard, which includes the input query format, output query format and query management tools related to the multimedia by content search, with the initial focus was on the searching of MPEG-7 based multimedia description data.

The main requirements for the standard include:

- **Input query format**: This should provide an ability to perform different query types including query by MPEG-7 and textual descriptions, free text queries, query by example, query by standardized unique identifier such as ISRC (International Standard Recording Code), query by spatio-temporal relationships, query by combined Boolean operations and empty queries for browsing purposes. Also required is the ability to communicate user preferences, usage history, result set preferences (sorting, grouping, size limits, format and structure).
- **Output query format**: This should specify a standard structure for the default result set, paging information for the result set, and an ability to express the match confidence of each result item.
- **Query management**: This includes service selection, exception messages, relevance feedback, and querying server capabilities.

To facilitate integration with the rest of the existing MPEG-7 standards, it is required that the M7QF standard adopt XML. Since XML is inherently human readable, using XML also provides intrinsic human readability of the queries formulated in M7QF.

IV. SUMMARY OF PROPOSALS SUBMITTED TO MPEG

**A. Technical University of Crete, Lab of Distributed Multimedia Information Systems & Applications (TUC/MUSIC)**

The MPEG-7 Query Language (MP7QL) [22] was proposed by TUC/MUSIC. The MP7QL has the MPEG-7 as the data model and allows for querying every aspect of an MPEG-7 multimedia content description. In addition, it allows the explicit specification of boolean operators and/or preference values in order to allow both the combination of the query conditions according to the user intentions and the expression of the importance of the individual conditions for the users.
The MP7QL allows combining all the elements and/or attributes present in the MPEG-7 descriptions, without any restrictions regarding the element order and the cardinality restrictions of the MPEG-7 types. This is achieved through re-engineered types that carry the semantics of the corresponding MPEG-7 types and have the following features: (a) they have no required attributes; (b) they allow any number of occurrences for their elements; and (c) all their sequences have been replaced by choices.

The MP7QL queries may utilize the users’ FASP and Usage History as context, thus allowing for personalized multimedia content retrieval. The FASP model supported is compatible with the MP7QL and has the model of the standard MPEG-7 FASPs as a special case. The MP7QL query results are represented as MPEG-7 documents, guaranteeing the closure of the results within the MPEG-7 space.

B. DMAG-UPC/UPF

This proposal was authored by DMAG-UPC/UPF and other partners of the VISNET-II European Network of Excellence (Telecom Italia Lab, EPFL and Warsaw University of Technology). Its main features were:

1. Embedded XQuery, allowing users to directly querying the XML MPEG-7 descriptors and descriptor schemes stored in a database with the W3C’s XML Query Language.
2. Specific constructors allowing Relevance feedback.
3. Specific constructors allowing searching within a previous result set.
4. An abbreviated condition expression language (for query-by-textual-description only) allowing to reduce the complexity of query instances for simple and frequently used conditions (e.g. (fileSize<1000) and (fileExtension="jpg")).
5. A default response format for the result set (specifying a default subset of metadata to be returned in absence of the specification of a format in the input query).
6. A Web Services binding (and the corresponding WSDL descriptors), and also an alternative HTTP binding which would allow embedding a query inside an URI. This URI-based invocation adheres to the SOAP 1.2 - Part 2 section 4.1 recommendations and follows the Representational State Transfer (REST) architectural style.

C. ETRI, KETI, Konkuk University

There were three proposals by Korean participants of ETRI, KETI and Konkuk University. One proposal specifies the overall structure and separation of the query format, with clear separation of input query format, output query format, and the query management tools. The second proposal specifies query condition constructs based on the ETSI 102 822-6-1 standard [23] and extended to support various complex algebraic operations and combinations of conditions. Furthermore, this proposal proposes a new construct (named indicator), which enables referencing various predefined conditions when building complex conditions, creating sub-queries that can be joined to provide full queries.

The third proposal describes a new requirement for asynchronous query modes. In the asynchronous mode, the client sends the request and immediately receives the acknowledgement from the server or the service provider with a response identifier and may disconnect the connection. Later in time, the client may come back to the server with the received response identifier to gather the prepared response. This is useful for complex queries, which would require a client to stay online for the duration of the query if answered synchronously.

D. NHK

The proposal submitted by NHK divides queries into two groups, one is for basic queries, which handle exact-matching retrieval and the other is for advanced queries, which handle other more complex queries such as similar-matching. For basic queries, XQuery combined with a Web Services interface was proposed. For advanced queries, a Web Services approach using pointing mechanisms was proposed; this approach does not limit the variety of retrieval services possible. The advantage of XQuery for MP7QF is the ability to provide a user with an easy-to-use, rich query language that can also be applied to metadata standards other than MPEG-7.

E. University of Wollongong

The University of Wollongong proposal (Multimedia Query Format - MQF) [24] focuses on simplicity, mainly influenced by SOAP [25] and SRU/SRW [26]. MQF was designed to act as a container format containing multimedia items similar to SOAP, with the idea of stacking the containers in a Reverse Polish Notation (RPN) sequence to provide the ability to form complex Boolean algebra queries.

Another key feature of the MQF proposal is the encapsulation of Boolean operators inside a special “operator” container, where some Boolean operators are predefined, but the operator format is left open for servers to implement their own operators.

F. University of Passau, Deutsche Telekom, Fraunhofer Ilmenau

The German proposal was submitted by members of the following three organizations: University of Passau, Deutsche Telekom and Fraunhofer Ilmenau. It considered all normative parts of the Call for proposal and recommended an XML schema containing types for a clear separation of input query format, output query format and management tools. The management tools consist of the one side of a framework specifying components for service management (how to search for and connect to MP7 databases), retrieval management (e.g., a validator, retrieval engine), query interpreter (responsible for interpreting a MP7QF query at the target database query system), etc. On the other side, the tools specified XML types for describing service capabilities, or user sessions.
A service in our nomenclature describes an MP7 database. The proposed input query format provided means for structuring a query in parts specifying the content and structure of the result set and in parts defining its condition part for filtering the data set. The content and structure of the desired result is formulated by a combination of URNs (based on a Classification Scheme) defining the respective MP7 D or DS and an XPath expression highlighting specific nodes (e.g., a title).

For the query condition, operators such as Boolean and Comparison as well as arithmetic operators were defined. Furthermore, the input query format implemented a referencing model for example media data and their descriptions in order to define them once and reference within query operations as often as necessary. The output query format provides types for describing result set items considering weighting and ranking functionality and for supporting grouping and paging.

V. THE MPEG-7 QUERY FORMAT

The resulting MPEG-7 Query Format (M7QF) draft standard was formed by combining the best features of each of the proposals described in Section IV. This comprises the following three main parts:

- Input query format.
- Output query format.
- Exceptions and server messages format.

Each of the three parts will be discussed further with some examples in Sections V-A, V-B and V-C, respectively.
An example of a query in M7QF is shown in Fig. 2, where the client is specifying a query by example with the example defined previously as resource named id12. An example of a resource declaration is shown in the example in Fig. 3.

Query conditions can be joined together to form a complex query by using Boolean operators in the form of AND, OR, NOT, and XOR. An example of a query using Boolean operators is shown in Fig. 4.

The Boolean functions of M7QF allow the use of weight to another resource by using the optional attribute preferenceValue. The example in Fig. 4 shows a query with two conditions: a free text query with weight value of 30 (denoted by its preferenceValue attribute) and another free text query with a weight value of 30. However, the second free text condition was modified with a Boolean NOT operator. The Boolean equivalent of the query shown in Fig. 4 is:

“Free Text 1” AND NOT “Free Text 2”

B. Output query format

As is the case with the input query format, the output format was also designed using each component’s proposed output format, where the best aspects of each proposal were combined to create the current draft version of the output format. The M7QF output query format was designed to be flexible with features including:

1. Paging ability: The output format is able to separate the result of a query into pages for easier viewing and/or handling.
2. Use of free text annotation in the replies: The server is able to send some free text annotation in the reply.
3. Inclusion of expiration date in the result set: The server is able to inform the client when the result of a query is no longer considered relevant.
4. Ability to use namespace other than MPEG7 in the output format: Flexibility and compatibility with other multimedia metadata standards.

The reply mainly consists of the following sections:

1. Global free text, where the server could supply a server-specific message regarding the output
2. Result items, which are the multimedia items existing in the server that satisfies the conditions given by the client
3. System messages, describing any errors, exceptions, warning or messages.

The output example in Fig. 5 shows a server result where the server replied that five multimedia resources that satisfies the result set is inside the Resource element, with attributes of the record number of the result relative to the result set, the rank of the result, and the confidence of how closely the item matches what the client wants. The resource item itself is present inside the ResultItem element in the form of Resource element. Note that in Fig. 5 the server also sends a free text annotation of the result set inside the GlobalFreeText element.

C. Exceptions and messages format

M7QF provides a facility for the server to communicate any errors or informational messages to the client encapsulated inside the SystemMessage element. The SystemMessage element can contain any status messages, warning messages and exception messages in any combination. An example of the use of the SystemMessage element is shown in Fig. 5, where the SystemMessage element contains a query successful message.

The general template of the messages format is the numeric code of the message along with a short textual description of the message.
VI. CONCLUSIONS

This paper has presented the work in progress towards the finalization of the MPEG-7 Query Format standard, including a description of the proposed technologies and the most recent version that has reached Committee Draft (CD) status. Ongoing work is focusing on important issues, which are not yet finalized, such as the Query Management Tools. A management layer is needed to allow the configuration of important operative parameters as the timeout, the operation mode or the server capabilities. Another important effort is taking place in order to extend the input format with more operators and query types. One of the goals is to allow the presence of more than one data flow at the same time, which will allow, for example, to define conditions involving more than one data set, like joins or unions. New operators, like those for manipulating arithmetic values, dates or character strings, are also being considered, along with the necessary extensions of the query algebra.

ACKNOWLEDGEMENT

[TUC/MUSIC] The work presented here was partially funded in the scope of the DELOS II Network of Excellence in Digital Libraries (IST – Project Record No 507618).

[DMAG] This work has been partly supported by the Spanish government (DMM-MM project, TSI 2005-05277) and the European Network of Excellence VISNET-II (IST-2005-2.41.6) under the European Commission IST 6th Framework Program.

REFERENCES