

2007

MQF: An XML Based Multimedia Query Format

K. Adistambha

University of Wollongong

C. H. Ritz

University of Wollongong, christian_ritz@uow.edu.au

I. S. Burnett

University of Wollongong, ian_burnett@uow.edu.au

Publication Details

This conference paper was originally published as Adistambha, K, Ritz, CH, Burnett, IS, MQF: An XML Based Multimedia Query Format, 2007 IEEE International Conference on Multimedia and Expo ICME 2007, Beijing, 2-5 July, 264-267.

MQF: An XML Based Multimedia Query Format

Abstract

MQF is a new XML based format designed to facilitate communication between disparate systems for applications involving multimedia query by content. Currently no standardized protocol exists which are able to provide flexibility in formulation of a query, such as the combination of any multimedia format (image, video, sound) to serve as query terms, combined with very complicated query conditions that can utilize a hierarchy of meta-search engines. In this work, we propose MQF as a flexible solution to serve as a communication format between a client and server for use in content based multimedia searching.

Disciplines

Physical Sciences and Mathematics

Publication Details

This conference paper was originally published as Adistambha, K, Ritz, CH, Burnett, IS, MQF: An XML Based Multimedia Query Format, 2007 IEEE International Conference on Multimedia and Expo ICME 2007, Beijing, 2-5 July, 264-267.

MQF: AN XML BASED MULTIMEDIA QUERY FORMAT

Kevin Adistambha, Christian H. Ritz and Ian S. Burnett

Whisper Labs, School of Electrical and Telecommunications Engineering
University of Wollongong, Australia

ABSTRACT

MQF is a new XML based format designed to facilitate communication between disparate systems for applications involving multimedia query by content. Currently no standardized protocol exists which are able to provide flexibility in formulation of a query, such as the combination of any multimedia format (image, video, sound) to serve as query terms, combined with very complicated query conditions that can utilize a hierarchy of meta-search engines. In this work, we propose MQF as a flexible solution to serve as a communication format between a client and server for use in content based multimedia searching.

1. INTRODUCTION

The immense popularity of video sharing web sites such as YouTube [1] and Google Video [2] illustrates the popularity of video media among users, but also highlight the popularity of sharing such user-created content with the world at large. The problem lies in describing the content that is hosted in the databases; MPEG has provided a multimedia content description standard in the form of MPEG-7 [3] which provides a standardized description scheme to annotate and describe the content of a multimedia data.

A universal format is needed to provide a language of communication between a client and a multimedia database, similar in function to the Structured Query Language (SQL) as used for textual databases. Such a format could then link disparate multimedia database systems regardless of their own technology for performing the pattern matching and decouples the client software from the actual database implementation. The scope of the problem addressed in this paper, therefore, is the communication between a client and a server as shown in Fig. 1. The goal is to allow multiple disparate multimedia database systems with different implementations to communicate in an unambiguous way with a client. The communication between systems as shown in Fig. 1 is an important basic building block for any multimedia search system.

Such a standard already exists in the area of library catalogue searching, in the form of Z39.50 [4] and

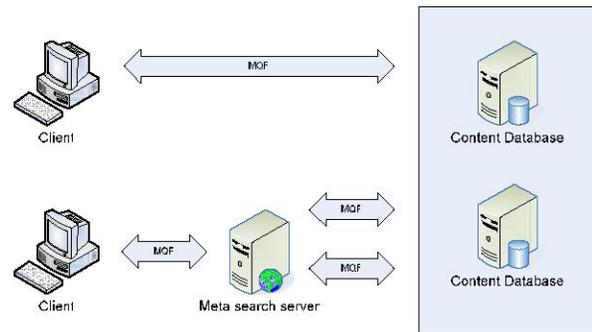


Figure 1. Scope of MQF.

Search/retrieve via URL (SRU) [5]. No such standard exists, however, in the area of multimedia content searching. For example, there is currently no standard way to search sites hosting video content other than by searching the textual metadata content, which is highly subjective and error prone since the metadata is entered by the user themselves. Since current technologies concerning content search are concentrated on image searching (e.g. MRML [6]), a new query format that is able to include video content searching using objective means is required.

This paper proposes a solution for a universal communication protocol between a client and a multimedia server, enabling the client to search for any multimedia content on a given server such as of sound, images, video or other media types. Section 2 will define the requirements and motivations for a modern multimedia query format, Section 3 will describe the proposed MQF, Section 4 will provide an explanation and an example instantiation of MQF, while Section 5 will present conclusions.

2. REQUIREMENTS AND MOTIVATIONS FOR A MODERN MULTIMEDIA QUERY FORMAT

2.1 Feature requirements

A new format has to be sufficiently flexible to allow future advancement in multimedia search technologies, but at the same time sufficiently strict as to not allow any ambiguity in the query formulated by the client. To anticipate the need to query multiple databases at once so that users are not restricted by a single search client, the protocol should also

```

<?xml version="1.0" encoding="UTF-8"?>
<mqf>

<queryId>ID_123</queryId>
<from>http://client.uri</from>
</from>
<query>
  <item queryLevel="example">
    <!-- Elements from Dublin Core or MPEG-7 -->
  </item>

  <operator>
    <!-- Operator -->
  </operator>

  <!-- Subsequent items & operators in RPN sequence -->
</query>
</mqf>

```

Figure 2. The general query input format of MQF.

be able to work in a meta-search engine environment. Therefore, a flexible multimedia query format should generally be:

- Human readable and easily understandable.
- Include methods to facilitate many query types e.g. query-by-example, exact search, query-by-metadata, etc.
- Language and character set independent.
- Able to use any multimedia format (image, sound, video, text) as a query term.
- Able to mix any media format, including textual formats, in a single query.
- Able to formulate very complex query conditions using simple operators.

An alternative to XML technology is the ISO/IEC 13249, also known as SQL Multimedia or SQL/MM [7]. However, historical implementation of the standard differs from system to system. Another notable drawback of SQL/MM is that the standard defines its own description scheme, rendering it incompatible with current multimedia description standards such as Dublin Core [8] and MPEG-7 [3]. An alternative is SRU, designed for searching library catalogues. However, the Common Query Language (CQL) [5], upon which it is based, is not designed for multimedia applications. As MRML [6] relies on the use of a finite state machine, this will not translate well into a multi-hierarchy meta-search engine environment, since all servers would have to keep track of the query state.

2.2 New generation search engines

The new format should enable the use of a meta-search server, where the meta-search server could keep a central ranking system, grading multimedia database servers by comparing the database servers by the relevancy of results it returns. In a meta-search engine, there exists a ranking of the

```

<?xml version="1.0" encoding="UTF-8"?>
<mqf>

<queryId>ID_123</queryId>
<from>http://server.uri</from>
</from>
<replies>
  <reply>
    <item>
      <!-- MPEG-7 MDS section 6.4 MediaLocator -->
    </item>

    <description>
      <!-- Any MPEG-7 description -->
      <!-- Any Dublin Core description -->
    </description>

    <index>
      <!-- Reply index -->
    </index>
  </reply>

  <!-- Subsequent reply items -->
</replies>
</mqf>

```

Figure 3. The general query output format/server reply of MQF.

relevancy of a database's response according to user supplied feedback. This relevancy ranking could be utilized for multiple databases hosting images, sound, video and textual data separately, grouped by media types. The database servers could then compete to provide the most relevant results.

2.3 Personal search engines

On the other hand, the new format also has to be sufficiently flexible to allow it to be used on a single server which acts as both a client and a server. A user could have a special lightweight search engine which only indexes and stores image files, for example. A client utilizing the new format then could be used to search the personal database of images, or search any server on the Internet that is compliant using the same interface.

3. MQF: AN XML BASED MULTIMEDIA QUERY FORMAT

To address the requirements and motivations described in Section 2, we have developed an XML based solution called the Multimedia Query Format (MQF). Key features of MQF are the use of XML so as to render it human and machine-readable, the ability to handle very complicated query conditions, the ability to work in a meta-search engine environment, the ability to work with any multimedia data, the use of a stateless design and non-specified client and server technologies for future-proofing.

3.1 Use of XML for human readability and character set independence

By using XML, the vast amount of tools available for XML parsing and development are immediately available for use by MQF. Use of the XML standard also allows MQF to be easily expandable to use any international standards for multimedia description that utilizes XML technology, such as MPEG-7 [3], MPEG-21 [9] and the Dublin Core [8] metadata initiative. Furthermore, any standard defined in MPEG-7/21 such as user preferences can be communicated to the server by using MQF as an envelope.

For presentation of results, a client or a server can transform the resulting XML formatted reply by using XSLT [10], hence minimizing the effort required for presentation. The use of XML technology directly provides MQF with language and character set independency, since XML was designed to be a general format for data interchange.

3.2 RPN notation for representing query conditions

In MQF, complicated query conditions such as “(A OR B) AND C” are efficiently communicated by utilizing Reverse Polish Notation (RPN). In RPN convention, the order of processing is implicitly stated and is from top to bottom; by using the RPN notation, exceedingly complicated queries can be created without ambiguity as to the order of operations. In contrast, traditional algebraic formulae must explicitly state the order of operations by means of brackets. RPN has been widely used in the Z39.50 standard [4].

By substituting numerical operands with any multimedia content as in MQF, any server is then able to perform very complex operations based on multimedia data unambiguously and without the need to signal which operand takes precedence over which. The server is able to parse the query input in sequence, hence minimizing the memory requirements while at the same time ensuring the accuracy of the output.

3.3 Query abstraction levels to remove ambiguity

To facilitate unambiguous interpretation of a query, the concept of query abstraction levels is introduced here.

There are four levels of query abstraction:

1. Exact level (“exact”): looking for an exact match; could be based on the descriptors provided. E.g. MPEG-7 descriptors, Dublin Core metadata.
2. Query by example level (“example”): looking for a similar match based on a given example.
3. Semantic description level (“semantic”): looking for a match by describing the semantics of the media. An example of a semantic description is Concept Languages [11].
4. Natural language level (“freetext”): looking for a match by vague, natural language description of an

event or a multimedia item

From the “exact” level to the “freetext” level, the query becomes more difficult to match exactly and requires more computation to process. The query abstraction concept is required to instruct the server how to interpret a certain query term and what kind of match is desired by the client. The abstraction level information is a required attribute for any query terms to avoid interpretation confusion by the server.

3.4 Flexible design to allow extensibility

The combination of query abstraction level and the use of RPN notation provides the ability of MQF to be as free form as possible, where a query can consist of multiple media formats, in different levels of abstraction and combined with a complicated query condition.

Note that this paper does not specify how a server should process a query except that the default response to the client should be in XML format. MQF is concerned only with how a client should formulate a query, resulting in flexibility on the server side. A server, for example, can choose not to process some query abstraction level due to processing power or other limitations. This flexibility allows advancement in technology to be implemented immediately on the server side without any impact on the communication protocol between the client and the server.

4. GENERAL ELEMENTS DESCRIPTION AND EXAMPLE INSTANTIATION

The general elements of MQF are outlined in Fig. 2:

1. The <queryId> and <from> elements identify the query relative to a client.
2. The <query> element in Fig. 2 signifies a query input operation, whereas <replies> in Fig. 3 signifies a server reply.
3. All query terms and operators are encapsulated inside the <query> element. Terms are in <item> element where the required attribute is the *queryLevel*, operators are in <operator> element. Possible values for the *queryLevel* attribute are “exact”, “example”, “semantic” and “freetext” as described in Section 3.3.
4. For transmitting binary data, base64 encoding specified by the MPEG-7 MDS Section 6.4 is used in MQF. This is shown in the example in Fig. 4.
5. The sequence of the <item> and <operator> elements in Fig. 2 is arranged according to the RPN sequence of terms.
6. All reply terms, including query terms, are encapsulated inside the <replies> element as individual <reply> elements which can contain a metadata description of each reply in both MPEG-7 and Dublin Core

```

<?xml version="1.0" encoding="UTF-8"?>
<mqf xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2001"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xsi:schemaLocation="urn:mpeg:mpeg7:mqf:schema:2007
mqf.xsd"
xmlns="urn:mpeg:mpeg7:mqf:schema:2007">
<queryId>ID_123</queryId>
<from>localhost</from>
<query>
<replyFormat>video/mpeg</replyFormat>
<item queryLevel="exact">
<dc:creator>John Doe</dc:creator>
</item>
<item queryLevel="example">
<mpeg7:InlineMedia type="image/jpeg">
<mpeg7:MediaData64>AaBbCcDd\=
</mpeg7:MediaData64>
</mpeg7:InlineMedia>
</item>
<operator>AND</operator>
</query>
</mqf>

```

Figure 4. An example of MQF query input instantiation.

elements.

The example in Fig. 4 shows an example instantiation of the query input format for searching for a video in MPEG format, created by John Doe, and includes an example JPEG image. The two terms ("John Doe" and the JPEG image) are connected with the Boolean AND operator in RPN sequence. The examples in Fig. 4 demonstrates the human readability aspect of MQF, where the RPN order of terms retains human readability in the XML instance in Fig. 4 regardless of the number of terms present in a query. The flexibility of formulating a query is also evident, due to the design of MQF to merely act as a container for query terms instead of predefining all of the allowed operation in a query. It is also relatively easy for a server to process the query, since the RPN order of terms enables the server to process the query in a stack-based fashion in sequence.

Fig. 5 shows the server reply, where the server specified that such an item exist in the URL <http://server/video1.mpg>, along with some Dublin Core and MPEG-7 elements describing the video. As in Fig. 4, the example in Fig. 5 shows that although the general format (as shown in Fig. 3) is relatively simple, the actual instantiation could contain many detailed information associated with the reply item.

5. CONCLUSION

This paper has introduced the Multimedia Query Format (MQF). This format is designed to serve as a format for querying multimedia databases by content. MQF makes it possible to formulate complex query conditions by using any combination of multimedia content in any format with a novel use of RPN notation and the introduction of the concept of Query Abstraction Levels. An example instantiation

```

<?xml version="1.0" encoding="UTF-8"?>
<mqf xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2001"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xsi:schemaLocation="urn:mpeg:mpeg7:mqf:schema:2007
mqf.xsd"
xmlns="urn:mpeg:mpeg7:mqf:schema:2007">
<queryId>ID_123</queryId>
<from>localhost</from>
<replies>
<reply>
<item>
<mpeg7:MediaUri>http://server/video1.mpg
</mpeg7:MediaUri>
</item>
<description>
<dc:title>Title 1</dc:title>
<dc:creator>John Doe</dc:creator>
<mpeg7:Classification>
<mpeg7:CaptionLanguage>en-US
</mpeg7:CaptionLanguage>
<mpeg7:ParentalGuidance>PG
</mpeg7:ParentalGuidance>
</mpeg7:Classification>
</description>
<index>1</index>
</reply>
</replies>
</mqf>

```

Figure 5. An example of MQF query output instantiation.

of MQF has highlighted its simplicity and flexibility. Its use of XML also allows MQF to work with existing web standards for information interchange such as SOAP and WSDL, and at the same time the abundance of XML tools are immediately available for MQF. MPEG has recognized the need for such a technology and has issued a call for proposal for the MPEG-7 Query Format; MQF was submitted to MPEG as a reply to the call for proposal.

6. REFERENCES

- [1] "YouTube: broadcast yourself," <http://www.youtube.com>.
- [2] "Google video," <http://video.google.com>.
- [3] ISO/IEC, "MPEG-7 Part 5: Multimedia Description Schemes (MDS)," ISO/IEC 15938-5:2001, 2001.
- [4] ANSI/NISO, "Z39.50-2003," <http://www.loc.gov/z3950/agency/>, 2003.
- [5] T. L. o. Congress, "Search/Retrieve via URL (SRU) Version 1.1," <http://www.loc.gov/standards/sru/>, 2004.
- [6] W. Müller, H. Müller, S. Marchand-Maillet, T. Pun, D. M. Squire, Z. Pečenović, C. Giess, and A. P. d. Vries, "MRML: A Communication Protocol for Content-based Image Retrieval," presented at Visual Information Systems, International Conference on, Lyon, France, 2000.
- [7] "SQL Multimedia and Application Packages," ISO/IEC JTC1/SC32/WG4, 2000.
- [8] "Dublin Core Metadata Initiative," <http://dublincore.org/>.
- [9] ISO/IEC, "MPEG-21 Multimedia Framework," ISO/IEC 21000.
- [10] W3C, "XSL Transformations (XSLT) Version 1.0," <http://www.w3.org/TR/xslt>, 1999.
- [11] J. A. Lay and G. Ling, "Semantic retrieval of multimedia by concept languages: treating semantic concepts like words," *IEEE Signal Processing Magazine*, vol. 23, pp. 123, 2006.