



Contents lists available at SciVerse ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

Uniform query formalization in mobile visual search: From standards to practice

Ruben Tous*, Jaime Delgado

Universitat Politècnica de Catalunya (UPC BarcelonaTech), Barcelona, Spain

ARTICLE INFO

Available online 17 February 2012

Keywords:

Standards
Visual search
Image retrieval
CBIR
Jpsearch
MPQF

ABSTRACT

Mobile visual search is the task of finding information (visual or not) given an input picture or video usually taken by a hand-held device. Interest around this topic is being boosted by the increasing amount of digitally stored images and the widespread proliferation of camera enabled mobile devices, such as mobile phones, PDAs or tablets. This fact is increasing the urgency of novel solutions for challenging problems such as the efficient coding of compact visual descriptors and the interoperability of distributed visual search query interfaces. Currently, almost every visual search service offers a different retrieval interface and image metadata description format, preventing unified and efficient access. In this context, standardization groups such as ISO/IEC SC29/WG11 (MPEG) and ISO/IEC SC29/WG1 (JPEG) have been working to create unified interfaces for image repositories. In one hand, MPEG provides the ISO/IEC 15 938-12 (MPEG Query Format, MPQF), which standardizes a query language for multimedia repositories and has also started an activity for standardizing compact descriptors for visual search (CDVS). On the other hand, JPEG is now finishing the ISO/IEC 24800 (JPSearch), which provides solutions to the image metadata interoperability problem. This paper analyzes how these standardization activities can be combined to satisfy the requirements posed by the mobile visual search scenario, which are their limitations and which would be the necessary actions to be taken by the standardization committees in order to overcome them.

© 2012 Published by Elsevier B.V.

1. Introduction

The impressive evolution in the generation and usage of digital images is stimulating the appearance of novel ways of searching and retrieving information, thus posing new challenges to the multimedia information retrieval community. A significant boost comes from the mobile visual search topic, due to the widespread proliferation of camera enabled mobile devices and the success of Web 2.0 applications. Mobile visual search allows overcoming the inherent limitations of text-based information retrieval systems by allowing to search and retrieve information (visual or not) from a hand-held device by just taking a photo or recording a video. Relying on text-based search interfaces and structured or unstructured text-based annotations is impractical for very large image databases, or for images that are generated continuously or ephemerally (e.g. from surveillance cameras, medical devices or smartphones). Due to its distributed nature, one of the problems to be addressed by practitioners and researchers of the mobile visual search topic is the one related to

* Corresponding author.

E-mail addresses: rtous@ac.upc.edu (R. Tous), jaime.delgado@ac.upc.edu (J. Delgado).

interoperability. There are already multiple existing mobile visual search systems but every one provides a different query interface and multimedia metadata description format. This fact prevents users from experiencing a unified access to the repositories. Systems aiming to provide a unified query interface to search images hosted in different systems without degrading query expressiveness need to address several questions which include but are not limited to the following:

- Which mechanism is going to be used for query formalization? How is this mechanism going to cope with visual search queries?
- Which metadata schema or schemas are going to be exposed to user queries? Is the system going to expose a mediated/pivot schema?
- How the mappings among the underlying target metadata schemas are going to be generated?
- Which formalism is going to be used to describe the mappings?
- How is the system going to use the mappings during querying?

Currently many standardization efforts are trying to provide answers to some of these questions. Two of the most relevant initiatives are the ISO/IEC 15938-12:2008 (MPEG Query Format or simply MPQF) [1–3] and ISO/IEC 24800 (JPEG’s JPSearch framework) [4–8]. While MPQF offers a solution for the query interface interoperability, JPSearch (whose Part 3 makes use of MPQF) faces the difficult challenge to provide an interoperable architecture for images’ metadata management. On the other hand, MPEG has also started an activity for standardizing compact descriptors for visual search (CDVS) [9]. This standard will unify the way to exchange image summaries in visual search queries. Developers of mobile visual search applications willing to provide a standard interface will face the problem of selecting the proper components of each one of these standards and try to figure out how to combine them. This paper¹ analyzes how these three standardization activities can be combined to satisfy the requirements posed by the mobile visual search scenario, which are their limitations and which are the necessary actions to be taken by the standardization committees in order to overcome them.

2. ISO/IEC 15938-12:2008 standard (MPEG query format)

A key element in all the different approaches to distributed image retrieval is the interchange of queries and API calls among all the involved parties. The usage of different proprietary interfaces for this task makes extremely difficult the deployment of distributed image search services without degrading the query expressiveness. The progressive adoption of a unified query interface would greatly alleviate this problem. In December 2008 the MPEG standardization committee (ISO/IEC JTC1/SC29/WG11) released a new standard which provides a standardized interface to search functionalities in distributed multimedia databases (including not only still images but also video and audio), the MPEG Query Format (MPQF). To achieve this goal, the MPQF standard specifies precise input and output parameters to express multimedia requests and uniform client side processing of result sets, respectively. Essentially, MPQF is an XML-based query language that defines the format of the queries and replies exchanged between clients and servers in a distributed multimedia search-and-retrieval system. In one hand, standardizing such kind of language fosters *interoperability* between parties in the multimedia value chain (e.g. content providers, aggregators and user agents). On the other hand, MPQF favors also *platform independence*; developers can write their applications involving multimedia queries independently of the system used, which fosters software reusability and maintainability.

2.1. MPEG query format syntax and terminology

MPQF queries (requests and responses) are XML documents that can be validated against the MPQF XML schema (see Fig. 1). MPQF instances include always the *MpegQuery* element as the root element. Below the root element, an MPQF instance includes the *Query* element or the *Management* element. MPQF instances with the *Query* element are the usual requests and responses of a digital content search process. The *Query* element can include the *Input* element or the *Output* element, depending if the document is a request or a response. The part of the language describing the contents of the *Input* element (requests) is named the Input Query Format (IQF) in the standard. The part of the language describing the *Output* element (responses) is named the Output Query Format (OQF) in the standard. IQF and OQF are just used to facilitate understanding, but do not have representation in the schema. Alternatively, below the root element, an MPQF document can include the *Management* element. Management messages (which in turn can be requests and responses) provide means for requesting service-level functionalities like interrogating the capabilities of a MPQF processor.

Example in Code 1 shows an input MPQF query asking for JPEG images related to the keyword “*Barcelona*”.

¹ This journal paper extends a manuscript appeared in the 7th International ICST Mobile Multimedia Communications Conference (MOBIMEDIA 2011), Cagliari, Italy, 2011 [10].

Download English Version:

<https://daneshyari.com/en/article/537721>

Download Persian Version:

<https://daneshyari.com/article/537721>

[Daneshyari.com](https://daneshyari.com)