A Search and Retrieval Framework for the management of copyrighted audiovisual content

Anna Carreras [1], Giovanni Cordara [2], Jaime Delgado [1], Frédéric Dufaux [3], Gianluca Francini [2], Thien M. Ha [3], Eva Rodríguez [1], Ruben Tous [1]

[1] Universitat Politècnica de Catalunya, Departament d'Arquitectura de Computadors, Campus Nord, Mòdul D6, Jordi Girona 1-3, E-08034 Barcelona, Spain

[2] Telecom Italia Lab, Via Reiss Romoli, 274 - 10148 Torino, Italy

[3] Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

E-mail: *annac@ac.upc.edu*

Abstract– *This paper presents a search and retrieval framework that enables the management of Intellectual Property in the World Wide Web. This twofold framework helps users to detect digital rights infringements of their copyrighted content. In order to detect possible copyright infringments, first the system crawls the Web to search replicas of user's images, and later evaluates if the copies have been taken respecting the terms stated by the owner. On the other hand, this framework also helps users in finding something interesting in the Web. It will provide copyrighted content to users according to their preferences and to intellectual property rights integrating search and retrieval with digital rights management tools.*

Keywords – *Content adaptation, duplicate detection, content search and retrieval, intellectual property rights*

1. INTRODUCTION

The number of web sites offering, copyrighted or free, multimedia content is continuously increasing. The ever growing volume of content requires effective search and retrieval systems, to help users in browsing and interrogating multimedia content repositories. Furthermore, there is no general solution dealing with copyright infringements, including adaptation or copies of original works.

This paper proposes a twofold framework. First, it enables the management of the Intellectual Property Rights related to audiovisual content in the World Wide Web. Second, the proposed framework offers an advanced search interface for querying audiovisual materials based on the novel MPEG Query Format (MPQF) [1] standard from the ISO/IEC JTC1/SC29/WG11 committee.

Two different applications are presented to show the feasibility of the proposed framework. On the one hand, the first application detects possible digital rights infringements for duplicate images. The system crawls the Web in order to find replicas of the users' images. Then, the duplicate detection service determines if images in the Web or in the Hidden Web, e. g. YouTube, are a duplicate of the original ones known to the system. If any duplicate is detected, then the Digital Rights Management (DRM) service determines if the copy has been taken according to the terms stated by the original creator. As well as detection of copyright-violation or other digital rights infringements, other possible applications are the monitoring of royalty payments, detection of illicit content, e.g. child pornography, or statistics.

On the other hand, a second application searches videos exploiting both textual descriptions and low level features that can be automatically extracted applying video analysis algorithms. These techniques enable the exploitation of all different modalities of interaction addressed by MPQF. Interrogating such multimedia content repository, users may express their preferences, e.g. the film genre, to shape the query results; the application presents to the user the items for which he has the appropriate licenses, or for which he can purchase the license.

2. BACKGROUND

2.1. MPEG Query Format

Existing metadata standards as the already completed ISO/IEC IS 15938, formally called the Multimedia Content Description Interface (better known as MPEG-7 [2]), provide rich models for multimedia content description. The usage of such standards should guarantee interoperability at data retrieval level. Generally, metadata descriptions link to standard schemas and should be interoperable with respect to XML query languages like W3C's XQuery [3].

However, data retrieval is not the only feature that a modern multimedia database should provide. Some content-based search and retrieval constraints must be also combined with the use of metadata descriptions. In this sense, XQuery does not provide specific normative functionalities for query-by-example or in general for content similarity measurement. A language which combines the expressive power of XQuery with content-based functionalities could be a good candidate for a unified query language for multimedia databases.

The MPQF is an XML-based language which defines the format of queries and replies to be interchanged between clients and servers in a multimedia information search and retrieval environment. MPQF allows combining Information Retrieval (IR) criteria with Data Retrieval (DR) criteria. MPQF can be used in standalone MultiMedia DataBases (MMDBs), but it has been specially designed for scenarios in which several MMDBs and content aggregators interact. MPQF does not make any assumptions about the metadata formats used by the target multimedia databases (MMDBs), which can be MPEG-7 but also any other format (Dublin Core for example).

2.1.1. MPQF overview

MPQF instances are XML documents that can be validated against the MPQF XML Schema. Any MPQF instance includes always the *MpegQuery* element as the root element. Below the root element, an MPQF document includes the *Query* element or the *Management* element.

MPQF instances with the *Query* element are the usual requests and responses of a multimedia search process. The *Query* element can include the *Input* element or the *Output* element, depending on whether the document is a request or a response. The part of the language describing the contents of the *Input* element is named the Input Query Format (IQF) in the MPQF standard. The part of the language describing the *Output* element is named the Output Query Format (OQF). Alternatively, below the root element, an MPQF document can include the *Management* element, which provides means for requesting service-level functionalities. The IQF is the part of MPQF which allows specifying the search condition tree and also the structure and desired contents of the server output. The condition tree is the main component of MPQF and can be built combining different kinds of expressions, including those related to the XML metadata with a complete set of operations related to content-based retrieval.

Within the IQF, but with respect to the output format specification, MPQF relies on XPath to offer flexibility when defining the structure of the XML result set. The OQF specifies which outputs from the server to the client are valid. The responses from an MPQF compliant server will always include a sequence of records (under the *ResultItem* tag), and within each record the client can find several MPQF standard fields like the URI of the related content or a text field, but also a complete or partial metadata description in an external format like MPEG-7.

2.2. Digital Rights Management Initiatives

DRM systems provide a means for content creators, distributors and rights issuers to address the unauthorized copying issue. A DRM system enables the management of multimedia content according to the permissions and constrains imposed by content creators, adapters and rights issuers.

Nowadays, there are several standard and commercial initiatives that specify a DRM system or the elements that form a DRM system and the relationships among them. Among the standards, the most relevant are the MPEG-21 standard [4] and OMA DRM [5]. Next section presents the MPEG-21 standard, since licenses governing content in the proposed framework are expressed using the Rights Expression Language defined in part 5 of this standard.

2.2.1. MPEG-21 standard

The MPEG-21 standard defines the elements needed to manage multimedia content through the digital value chain. It is divided in several parts that deal with different aspects of multimedia information management.

Part 5 of the MPEG-21 standard [6] defines a Rights Expression Language (REL) for declaring permissions and constraint of use of digital content. The most important concept of the MPEG-21 REL are licenses that can be conceptualised as containers of grants, each grant of which conveys to a principal or entity the right of exercising an action against a resource if some conditions are met. Another important concept of the MPEG-21 REL is the authorisation model, which can be used by any software implementation that makes an authorisation decision using MPEG-21 REL licenses. The central question of this decision making process is: *"Is a principal authorised to exercise a right against a resource?"*.

In DRM systems compliant with the MPEG-21 standard, licenses are generated according to the MPEG-21 REL containing the content usage rules and they are associated to an entity and to a digital resource. Then, they can be enforced by the DRM system. Licenses can be associated with the digital resources that they govern in four different ways: the license could be embedded in the Digital Item that contains the resource; the license can be referenced within the Digital Item; the Digital Item can refer to a license service; or the license can refer to the resource that it is governing.

3. FRAMEWORK FOR DUPLICATE IMAGE AND DIGITAL RIGHTS INFRINGEMENTS DETECTION

The framework for duplicate image and digital rights infringement crawls the web to search duplicates of user images and determines if copying has been done respecting intellectual property rights of the original ones. The image duplicate detection service aims at recognizing modified versions of a

given original image after common image processing operations, as long as the latter do not alter the perceptual meaning of the image content. The DRM service detects copyright infringements in the replicas of the users' images found in the Web and Hidden Web. Next sections present in detail the different modules that make up the framework.

3.1. Multimedia content retrieval service

The framework exposes multimedia search and retrieval functionalities with an MPQF interface. MPQF provides the ability to extend access to the metadata and contents via a standard query interface, in the same way as Z39.50, but making use of the newest XML querying tools (based on XPath 2.0 and XQuery 1.0) in combination with a set of advanced multimedia information retrieval capabilities defined within MPEG. With respect to still and moving images, MPQF provides several content-based search functions like QueryByMedia (query by example media), QueryByROI (query by media with region of interest), QueryByDescription (query by example description), QueryByFeatureRange, SpatialQuery and TemporalQuery. A proper combination of these functions allows fulfilling a broad range of potential user information needs. A special interest for this work has the QueryByMedia function, which can be used as an interface for searching image replicas within a database.

3.2 Duplicate detection service

The image duplicate detection service is composed of two components: a set of binary detectors trained on each given original image, and an indexing mechanism to quickly identify the set of most likely candidates [7].
More specifically, each binary detector is trained on a given original image, and determines whether a test image is a duplicate of this original. For this purpose, the following visual features are used: texture using statistics of Gabor transform, color classes in the Hue Saturation Intensity (HSI) space, and grey-level classes based on the equalized intensity channel. Statistical normalization is performed to give the same significance to each feature, and the space dimensionality is then reduced by means of Principal Component Analysis (PCA). Finally, the detector determines the probability that the test image is a duplicate of the given original image by means of a decision function based on support vector classifier.
In order to extend this detector to a collection of original images, an indexing technique is first used. It identifies a set of most likely originals, or candidates, corresponding to a test image. The above described binary detectors are then applied for each candidate, returning the best match. Finally, the

duplicate detection service decides whether the test image is a duplicate of this original if the corresponding probability is above a threshold.

3.3. DRM service

The DRM service detects copyright-violation or other digital rights infringements for still and moving images. It provides functionality to obtain the licenses governing a digital resource, in this case an image, and to verify if copies have been done according to the rights and conditions stated by the creator of the original one. More specifically, it offers the following two operations:
- *getLicenses*: It receives as parameters an image label and returns an XML file containing the set of MPEG-21 REL licenses governing the image
- *verifyRights*: It receives as parameters the licenses governing a duplicated image and the XML file containing information about the usage, and determines if the terms stated by the image creator has been respected in the different replicas of the image. This operation implements a license verification algorithm, based on the MPEG-21 REL Authorization Model [6], which verifies if an entity was authorised to perform a replica of the original image.

4. VIDEO SEARCH FRAMEWORK

Addressing an advanced web application that provides copyrighted multimedia content, one has to face the problem of providing ways to help users in finding something interesting with new, effective and appealing methods. Furthermore, the retrieved results need to be supplied with the associated licenses or with convenient ways for pursuing them. This results in the need of a complete platform that integrates DRM tools with query interface and search and retrieval engine. Therefore, a flexible application enabling all the different queries addressed by MPQF and, consequently, algorithms providing answers in line with users' expectations are required. The presented video search framework tries to address these problems, representing a search and retrieval service that integrates a search and retrieval engine, and DRM management tools. The technologies related to these two aspects are described in deeper details in the following subsections.

4.1 Audiovisual content search and retrieval service

A search and retrieval service able to respond to generic queries expressed with MPQF needs to be integrated with effective search and retrieval algorithms.

A technology that has proven to guarantee good performances in different use cases is the analysis of textual metadata (keywords, textual descriptions, plots, actors, user comments ...). There are several standards (like MPEG-7 [2]) describing information related to multimedia contents in a textual and interoperable format. In a real scenario, however, video content is not always accompanied by other corresponding information. There is the need, therefore, to provide innovative ways to allow users to search for content exploiting all the available information. A solution can be identified in the automatic analysis of visual information. Nevertheless, one has to face the problem of getting over the semantic gap, i.e. the gap between low level features (LLFs) that can be automatically extracted from a digital content and the interpretation that a user would have of the same content. Of course the problem of closing the semantic gap is difficult to solve since it is intrinsically embedded in the nature of video digital contents and strictly related to human interpretation. A way to address the problem and improve the overall performances that is being explored inside VISNET II [8] is to analyze LLF jointly with textual metadata, whenever available.

The technical work conducted in this activity can be described as a sequence of different operations:

- Automatic extraction of low level features: for each video, temporal and spatial features are extracted, the former related to the temporal evolution of the videos, revealing therefore aspects of storytelling style; the latter, represented by a set of MPEG-7 descriptors, provides information about visual aspects;

- Analysis of textual information: Latent Semantic Indexing (LSI) technique has been used [9], a vector space technique that exploit co-occurrences between terms. Using LSI it is possible to discover similarities between texts even if they share few or no words;

- Construction of searchable indexes: The data extracted with textual and visual analysis are jointly used for creating tables of distances between contents in the repository. Such tables can be used in real time to provide answers to different kind of queries.

For further information one can refer to [10].

4.2 DRM service

The DRM service determines if the user can use the videos found by the audiovisual search and retrieval service. It provides functionalities to obtain the licenses governing a digital resource, in this case a video, and provides information on content usage, according to licenses governing the selected video, to the user. This service provides with two operations. The first one obtains the licenses associated to the video selected by the user, as in the Framework for duplicate image and digital rights infringements detection. The second one determines the user's permissions and constraints of content usage.

5. CONCLUSION

A search and retrieval framework that enables the management of digital rights was presented. This framework offers an advanced search interface for querying copyrighted digital media.

Two different applications have been presented to demonstrate the feasibility of the proposed twofold framework. The framework enables users to detect possible digital rights infringements of replicas of user's images in the Web or in the Hidden Web, e. g. YouTube. The framework also enables users to find interesting copyrighted content in the Web by means of an advanced search interface for querying audiovisual materials.

REFERENCES

[1] ISO/IEC/SC29/WG11/N9341. *"ISO/IEC 15938-12 FCD MPEG Query Format"*, October 2007

[2] http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm

[3] XQuery 1.0: An XML Query Language. W3C Proposed Recommendation 21 November 2006. See http:/www.w3.orgTR/xquery/.

[4] http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm

[5] OMA DRM Specification, *"OMA-TS-DRM-DRM-V2_0-20050901-C"*. September 2005.

[6] ISO/IEC, *"ISO/IEC IS 21000-5 – Rights Expression Language"*. April 2004

[7] Y. Maret, D. Marimon Sanjuan, F. Dufaux and T. Ebrahimi, *"Hierarchical Indexing using R-trees for Replica Detection"*, in SPIE Proc. Applications of Digital Image Processing XXIX, San Diego, CA, August 2006

[8] VISNET II, http://www.visnet-noe.org/index.html

[9] M.W. Berry, S.T. Dumais, and G.W. O'Brien. *"Using Linear Algebra for Intelligent Information Retrieval"*, SIAM Review, Vol. 37, No. 4, Dec. 1995, pp. 573-595..

[10] IST-1-038398 - Networked Audiovisual Media Technologies - VISNET II, "Deliverable D2.2.5: First set of developments and evaluation for search systems for distributed and large audiovisual databases". November 2007.