An Infrastructure for MultiMedia Metadata Management

Patrizia Asirelli, Massimo Martinelli, Ovidio Salvetti

Istituto di Scienza e Tecnologie dell'Informazione, CNR,
56124 Pisa, Italy
{Patrizia.Asirelli, Massimo.Martinelli,
Ovidio.Salvetti}@isti.cnr.it
http://si-lab.isti.cnr.it

Abstract. This paper presents an approach for the integration of multimedia metadata and their management based on Semantic Web technology. In particular, we propose a java-based Infrastructure for MultiMedia Metadata Management – 4M - composed of five main components, an *MPEG-7 feature processing* unit, an *XML database management* unit, an *algorithms ontology-exploiting* unit, a *multimedia semantic annotation* and *integration* units. This way, we intend to introduce the novel idea of managing also algorithms on a variety of multimedia metadata (audio, images and videos) to add the capability of tracking data processing. This work is mainly carried out in the framework of the European Network of Excellence MUSCLE (Multimedia Understanding through Semantics, Computation and Learning), where ISTI-CNR is leading the 'Representation and Communication of Data and Metadata' Workpackage.

1 Introduction

Nowadays, a huge amount of multimedia information produced by an increasing number of connected digital devices is becoming available on the Web. Obviously, one can imagine that this amount will only be growing in the future.

Multimedia management on the Web is a hot topic and many research teams, projects and working groups are active in this area: W3C [1], DELOS [2], aceMedia [3], etc..

MUSCLE [4] aims at supporting a research Network of Excellence (NoE) to foster close collaboration between research groups in multimedia data mining and machine learning. Within MUSCLE we are working to show possible strategies for the interoperability of multimedia groups. In particular, the WP9 Workpackage ('Representation and Communication of Data and Metadata' [5]) is dedicated to enabling interaction and exchange of metadata emanating from different multimedia modalities. Thus, our activity primarily concerns the definition of a strategy for the NoE to develop, maintain, and provide access to multimedia metadata and data sets. It aims at providing an integrated metadata environment able to support eventually different metadata standards and uses (browsing, search, translations, etc.).

Additionally, as "Signals and Images" group we have years of experience within the field of image representation and understanding. We have been involved in many projects (e.g., ESPRIT, BRITE, CRAFT) and applications areas, such as health care, industrial quality control, cultural heritage, and others. Thus, we are very much aware of the importance of instruments that allow the exchange, retrieval and elaboration of multimedia data and metadata on the Web. Besides, we acknowledge the need to add semantic annotation to image representation as a value-added contribution to possible applications that require intelligent, non-naive retrieval and possibly "reasoning". Within collaboration with the Russian Academy of Sciences (RAS) we are in fact considering the implementation of an ontology for image understanding in line with the latest Semantic Web technology [6].

Several research groups are working in the field of multimedia management (DELOS [2], aceMedia [3], …, IBM [7]) and the results are not stable yet. A comprehensive solution still does not exist even if many specific solutions are being studied and/or developed.

Some standards have been proposed, in particular MPEG-7 [8] is the most mature and stable. Unfortunately, while MPEG-7 compliant descriptions can be easily created, the freedom in terms of structures and parameters is such that generically understanding MPEG-7 produced by others is quite difficult.

The use of technologies coming from Semantic Web, promoted by the W3C, could facilitate the overall vision of distributed, machine readable metadata on Internet. To enable this scenario, standardized frameworks have been developed to express semantic relationships between resources (RDF [9]), ontologies describing domain classes and their relationships (RDFS [10] and OWL [11]).

In this paper, based on the ideas presented at EWIMT 2005 [12], we propose a java-based infrastructure using most mature open-source tools and standards. This infrastructure is composed of five main components, an *MPEG-7 feature processing* unit, an *XML database management* unit, an *algorithms ontology-exploiting* unit, *multimedia semantic annotation* and *integration* units. This way, we intend to introduce the novel idea of managing also algorithms on a variety of multimedia metadata (audio, images and videos) to add the capability of tracking data processing.

2 Aim and scope

The infrastructure for multimedia metadata management could be used in various contexts and by users with different profiles.

Among the possible scenarios, we focused mainly on two use cases:

> a. a user needs to manage a collection of personal information (e.g., photos, videos, music, and so on) and then retrieve specific items under certain semantic conditions (e.g., photos showing smiling persons);

b. users in a network need to share multimedia resources and related semantic information: some resources should be private and accessible only by the owners, while some others should be public or shareable (sub-network).

Consequently, the system should provide the following capabilities:
- to store, organize and retrieve distributed multimedia resources;
- to manage algorithms for information processing;
- to add semantic annotations;
- to access and share information.

Furthermore, a public provider could permit to access services (servers, grid, ...) or else users could peer-to-peer connect to others. Finally, semi authoritative models (as described in [12]) could be taken into account.

3 Overview of the infrastructure characteristics
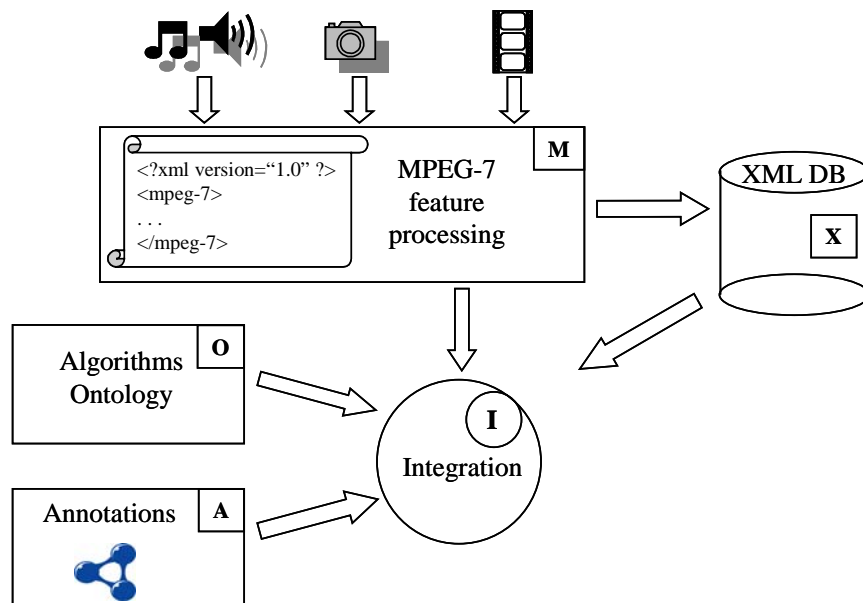
Figure 1 shows the architecture we are working on:


```
<?xml version="1.0" ?>
<mpeg-7>
. . .
</mpeg-7>
```

Fig. 1. 4M Infrastructure.

- Unit "M" is devoted to MPEG-7 features processing from multimedia objects;

- Unit "X", based on an XML [13] database, manages MPEG-7 features organized as XML files;
- Unit "O" is based on an ontology of algorithms describing processes and procedures applied to multimedia objects;
- Unit "A" adds annotations to multimedia objects to describe specific semantic information;
- Unit "I" provides interfaces and tools to integrate and access the other modules.

- MPEG-7 Features Processing Unit

Currently, XML is the only standard format officially approved for MPEG-7 (XML Schema). In fact, some proposals have been defined in OWL [14, 15, 16] but still it is not clear which one will emerge.

Only few programs, among those allowing MPEG-7 feature extraction, are completely open-source. In particular, we have considered "XM - eXperimental Model" [17], written in C++, Mpeg7Audio Encoder [18] and Vizir [19], both written in Java.

We decided to consider MPEG-7AudioEncoder and Vizir codes. We improved them by introducing few extensions to the code in order to extract MPEG-7 features from audio and still images.

Currently, the system we have developed is able to extract almost all MPEG-7 features from audio, color and texture from still-images, and a restricted set of features from videos.

- XML Database Management Unit

Several solutions have been investigated. Discarding commercial ones, we have examined only open-source projects.

All the solutions have pros and cons, since no one reaches the expected requirements. As a first choice we have then decided to consider only four of them (Berkeley DB XML, eXist, Ozone XML and Xindice) and, finally, we have chosen "eXist" [20], that won the Infoworld 2006 Technology of the year award. It allows the use of the XQuery [21] language to retrieve information and it provides Java APIs to manage the DB. Anyway, there are still problems (e.g. at the moment of writing this paper, this product doesn't allow concurrency, but the new beta version seems to do it).

Given that an URI (Unique Resource Identifier) is the basic brick for Semantic Web applications, we have thus decided to denote any multimedia resource by a unique identifier, a MediaURI, which includes the type of object and a hash of the content of the object. Through a MediaURI, any multimedia resource is stored and identified in the XML database.

- Algorithms Ontology-Exploiting Unit

An ontology of algorithms should be used to describe processes needed to obtain multimedia objects. Currently, in a collaborative project with the Dorodnicyn Computing Center of the Russian Academy of Science, an ontology of image algorithms is under development in OWL. It is based on a technical vocabulary of more than 1000 terms describing image characteristics, algorithms used to obtain images, and relations among terms.

- Multimedia Semantic Annotations Unit

In order to add semantics describing information inherent to particular application areas (e.g. health care information, personal photo DB, etc...), a tool for annotating multimedia objects is under investigation. At present, the only annotations that we are handling are based on MPEG-7. We intend to overcome the limits of this approach by introducing rules in order to improve searching and reasoning capability (e.g. using SWRL [22]).

- Integration Unit

All units previously discussed should interact and be accessed and managed by the Integration Unit to allow retrieving and adding the required information through suitable tools and interfaces.

Presently, we are investigating inference engines based on OWL and SWRL, and java tools such as Jena [23], Jess [24] and Mandarax [25].

Let us point out that, in any case, all units are independent so that the infrastructure is scalable, i.e. it should allow new technologies to be easily integrated.

4   Conclusions

An approach for the integration of multimedia metadata and their management based on Semantic Web technology has been discussed. It consists of a java-based Infrastructure for MultiMedia Metadata Management - 4M - composed of five main components, an *MPEG-7 feature processing* unit, an *XML database management* unit, an *algorithms ontology-exploiting* unit, a *multimedia semantic annotation* and *integration* units. This way, we intend to introduce the novel idea of managing also algorithms on a variety of multimedia metadata (audio, images and videos) to add the capability of tracking data processing. To implement this infrastructure, the state of the art in the Semantic Web area has been considered and, in particular, some opensource tools have been selected. This work is mainly being carried out in the frame of the MUSCLE-NoE, WP9 workpackage, and then it describes an ongoing activity. Part of this infrastructure has already been implemented. Future work will focus on the completion of the infrastructure itself and its experimentation within the network.

Acknowledgements

References

1. W3C World Wide Web Consortium http://www.w3.org/
2. DELOS Network of Excellence http://www.delos.info//
3. aceMedia project http://www.acemedia.org/aceMedia/
4. MUSCLE Network of Excellence 'Multimedia Understanding through Semantics, Computation and Learning' http://www.muscle-noe.org/
5. MUSCLE Workpackage 9 http://muscle.isti.cnr.it/
6. Beloozerov, V.N., Gurevich, I. B., Gurevich,N. G., Murashov, D.M., Trusova, Y.O.: Thesaurus for Image Analysis: Basic Version, Pattern Recognition and Image Analysis, Vol. 13, No. 4, 556–569 (2003).
7. IBM http://www.ibm.com
8. MPEG-7 http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm
9. RDF http://www.w3.org/RDF/
10. RDF Schema http://www.w3.org/TR/rdf-schema/
11. OWL http://www.w3.org/2004/OWL/
12. Asirelli, P., Di Bono, M.G., Martinelli, M., Salvetti, O., Signore, O., Catasta, M., Morbidoni, C., Piazza, F., Tummarello, G.: Toward a Scalable Multimedia Metadata Infrastructure using Distributed Computing and Semantic Web Technologies. IEE Proc. EWIMT (2005) 153-156
13. XML eXtensible Markup Language http://www.w3.org/XML/
14. Hunter, J.: Adding Multimedia to the Semantic Web - Building and Applying an MPEG-7 Ontology. Multimedia Content and the Semantic Web: Standards, Methods and Tools, Giorgos Stamou and Stefanos Kollias (Editors), Wiley (2005).
15. Garcia, R., Celma, O.: A Complete MPEG-7 OWL Ontology based on a XMLSchema to OWL Mapping, www.acemedia.org/aceMedia/files/multimedia_ontology/presentations_1st_meeting/special_event_EWIMT05-celma.pptMPEG-7 Ontos http://dmag.upf.edu/ontologies/mpeg7ontos/
17. XM experimental Model http://www.lis.ei.tum.de/research/bv/topics/mmdb/e_mpeg7.html
18. Mpeg7Audio Encoder http://www.ient.rwth-aachen.de/team/crysandt/software/mpeg7audioenc/
19. Vizir http://vizir.ims.tuwien.ac.at
20. eXist http://exist.sourceforge.net/
21. XQuery http://www.w3.org/XQuery/
22. SWRL http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/
23. Jena http://jena.sourceforge.net/
24. Jess http://herzberg.ca.sandia.gov/jess/
25. Mandarax http://mandarax.sourceforge.net/