O

Ontology and Multimedia

Roberto Poli
University of Trento, Italy

Achilles Kameas
Hellenic Open University, Italy

Lambrini Seremeti
Hellenic Open University, Italy

INTRODUCTION

Audiovisual resources in the form of still pictures, graphical, 3D models, audio, speech, and video play an increasing pervasive role in our lives, and there will be a growing need to manage all these multimedia objects. This is a task of increasing importance for users who need to archive, organize, and search their multimedia collections in an appropriate fashion.

To cope with this situation, much effort has been put in developing standards both for multimedia data (natural and synthetic (e.g., photography, face animation), continuous and static (e.g., video, image)) and for data describing multimedia content (metadata). The aim is to describe open multimedia frameworks and achieve a reasonable and interoperable use of multimedia data in a distributed environment.

BACKGROUND

Metadata are a representation of the administrative, descriptive, preservation, usage, and technical characteristics associated with multimedia objects; they can be extracted manually or automatically from multimedia documents. This value-added information helps bridge the semantic gap, described as: "The lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation" (Smeulders, Worring, Santini, Gupta, & Jain, 2000).

Because of the high cost and subjectivity associated with human-generated metadata, a large number of research initiatives are focusing on technologies to enable automatic classification and segmentation of digital resources. Many consortia are working on a number of projects in order to define multimedia metadata standards, which are being developed in order to describe multimedia contents in many different domains and to support sharing, exchanging, and interoperability across different networks. They are distinguished in Salvetti, Pieri, & Di Bono, 2004):

- *Standardised description schemes* that are directly related to the representation of multimedia content for a specific domain (like METS, MPEG-7).
- *Standardised metadata frameworks* that consider the possibility of integrating more metadata standards mapped on different application domains, providing rich metadata models for media descriptions together with languages allowing one to define other description schemes for arbitrary domains (like PICS, RDF, MPEG-21).

For example, the vision of MPEG-21 is to define a multimedia framework to enable augmented and transparent use of multimedia resources across a wide range of networks and devices used by different communities. The intent is that this framework will cover the entire multimedia content delivery chain, including creation, production, delivery, personalization, presentation, and trade.

The development of metadata standards will increase the value of multimedia data, which are used by various applications. Nevertheless, there are disadvantages in current metadata representation schemes (Smith & Schirling, 2006). Some of them are cost, unreliability, subjectivity, lack of authentication, and interoperability with respect to syntax, semantics, vocabularies, and languages (Salvetti et al., 2004).

It is necessary to have a common understanding of the semantic relationships between metadata terms from different domains. Representation and semantic annotation of multimedia content have been identified as an important step toward more efficient manipulation and retrieval of multimedia. In order to achieve semantic analysis of multimedia content, ontologies are essential to express semantics in a formal machine-processable representation (Staab & Studer, 2004).

Professional groups increasingly are building metadata vocabularies (or ontologies). A number of research and standards groups are working on the development of common conceptual models (or upper ontologies) to facilitate interoperability between metadata vocabularies and the integration of information from different domains.

MAIN FOCUS OF THE ARTICLE

Multimedia Ontologies

Ontologies have applications in many areas, including natural language translation, medicine, standardization of product knowledge, electronic commerce, and geographic information systems, among others. Many of these applications use or will use multimedia data in the immediate future, making the creation of multimedia ontologies a crucial component (Alejandro & Smith, 2003).

It is well known that the word "ontology" generates a lot of controversy in discussions about Artificial Intelligence, although it has a long history, in which it refers to the categorical framing of what is (Poli, 2001, 2002, 2007; Poli & Simons, 1996). Briefly it can be claimed that ontology deals with what can be rationally understood, at least partially. According to this interpretation, science in all of its branches is the most successful and powerful ally of ontology.

One may say that there are material things, plants and animals, as well as the products of the talents and activities of animals and humans in the world. This first almost trivial list already indicates that the world comprises not only things, animate or inanimate, but also activities and processes and the products that derive from them. It is likewise difficult to deny that there are thoughts, sensations, and decisions, and the entire spectrum of mental activities, just as one is compelled to admit that there are laws and rules, languages, societies, and customs. We can set about organizing this list of objects by saying that there are independent items that may be real (mountains, flowers, animals, and tables), or ideal (sets, propositions, values), and dependent items which in turn may be real (colors, kisses, handshakes, and falls) or ideal (formal properties and relations). All these are in various respects items that are.

A multimedia ontology, informally, is a means for specifying the knowledge of the world through multimedia documents in a structured way such that users and applications can process the descriptions with reference to a common understanding.

Four different levels of information are represented in multimedia ontologies (Euzenat et al., 2004):

- Signal information
- Featural information
- Symbolic information
- Semantic information

Multimedia ontologies can be of two types:

- Media-specific ontologies have taxonomies of different media types and describe properties of different media. For example, video may include properties to identify the length of the clip and scene breaks.
- Content-specific ontologies describe the subject of the resource, such as the setting or participants. Because such ontologies are not specific to the media, they could be reused by other documents that deal with the same domain.

Multimedia ontologies are used for different goals by different applications, including the following (Alejandro & Smith, 2003):

- **Content visualization:** They can be used to create tables of content and used for browsing.
- **Content indexing:** They can be used to improve indexing consistency in manual annotation systems (Schreiber et al., 2001) (e.g., use the term apartment instead of flat), or in the propagation of labels in automatic indexing systems (e.g., a face detected implies a person was detected).
- **Knowledge sharing:** Annotated multimedia collections can be more easily shared if they use a common conceptual representation.

Related Content

Adapting Multimedia Streaming to Changing Network Conditions
Dimitris Kanellopoulos (2011). *Streaming Media Architectures, Techniques, and Applications: Recent Advances (pp. 24-38).*
www.irma-international.org/chapter/adapting-multimedia-streaming-changing-network/47513

Adapting Web Sites for Mobile Devices - A Comparison of Different Approaches
Henrik Stormer (2006). *Handbook of Research on Mobile Multimedia (pp. 278-290).*
www.irma-international.org/chapter/adapting-web-sites-mobile-devices/20971

Introduction to Multicast Technology
Gabor Hosszu (2002). *Multimedia Networking: Technology, Management and Applications (pp. 369-411).*
www.irma-international.org/chapter/introduction-multicast-technology/27041

Towards Fusion of Textual and Visual Modalities for Describing Audiovisual Documents
Manel Fourati, Anis Jedidi, Hanen Ben Hassinand Faiez Gargouri (2015). *International Journal of Multimedia Data Engineering and Management (pp. 52-70).*
www.irma-international.org/article/towards-fusion-of-textual-and-visual-modalities-for-describing-audiovisual-documents/130339

An Automatic Video Reinforcing System for TV Programs using Semantic Metadata from Closed Captions
Yuanyuan Wang, Daisuke Kitayama, Yukiko Kawai, Kazutoshi Sumiyaand Yoshiharu Ishikawa (2016). *International Journal of Multimedia Data Engineering and Management (pp. 1-21).*
www.irma-international.org/article/an-automatic-video-reinforcing-system-for-tv-programs-using-semantic-metadata-from-closed-captions/149229