INTEGRATING KNOWLEDGE, SEMANTICS AND CONTENT FOR USER-CENTRED INTELLIGENT MEDIA SERVICES: THE ACEMEDIA PROJECT

I. Kompatsiaris[2]*, Y. Avrithis*[2]*, P. Hobson*[1] *and M. G. Strintzis*[2]

[1] Motorola Ltd
Personalisation and Knowledge Lab
Jays Close, Basingstoke RG22 4PD, UK
Email: Paola.Hobson@motorola.com

[2]Informatics and Telematics Institute
1st Km Thermi-Panorama Rd
Thessaloniki 57001, Greece
Email: {ikom, iavr}@iti.gr

ABSTRACT

In this paper, an approach for knowledge and context-assisted content analysis and reasoning based on a multimedia ontology infrastructure is presented. This is one of the major objectives of the aceMedia Integrated Project. In aceMedia, ontologies will be extended and enriched to include low-level audiovisual features, descriptors and behavioural models in order to support automatic content annotation. This approach is part of an integrated framework consisting of: user-oriented design, knowledge-driven content processing and distributed system architecture. The overall objective of aceMedia is the implementation of a novel concept for unified media representation: the *Autonomous Content Entity* (ACE), which has three layers: content, its associated metadata, and an intelligence layer. The ACE concept will be verified by two user focused application prototypes, enabled for both home network and mobile communication environments.

1. INTRODUCTION

Video understanding and semantic information extraction have been identified as important steps towards more efficient manipulation of visual media. Although new multimedia standards, such as MPEG-4 and MPEG-7, provide the needed functionalities in order to manipulate and transmit objects and metadata, their extraction, and that most importantly at a semantic level, is out of the scope of the standards and is left to the content developer.

In well-structured specific domain applications (e.g. sports and news broadcasting) domain-specific features that facilitate the modelling of higher level semantics can be extracted [3], [4]. A priori knowledge representation models are used as a knowledge base that assists semantic-based classification and clustering [5, 6]. In [7], semantic entities, in the context of the MPEG-7 standard, are used for knowledge-assisted video analysis and object detection, thus allowing for semantic-level indexing. In [8], fuzzy ontological relations and context-aware fuzzy hierarchical clustering are employed to interpret multimedia content for the purpose of automatic thematic categorization of multimedia documents. In [9] the problem of bridging the gap between low-level representation and high-level semantics is formulated as a probabilistic pattern recognition problem. Finally, in [10] and [11], hybrid methods extending the query-by-example strategy are developed.

Due to the limitations of the state of the art multimedia analysis systems [12], it is acknowledged that in order to achieve semantic analysis and knowledge mining from multimedia content, ontologies are essential to express the key entities and relationships describing multimedia in a formal machine-processable representation [13]. Ontology modelling and ontology-based metadata creation currently address mainly textual resources [14] or simple annotation of photographs [15]. aceMedia will investigate ontology modelling in terms of both methodology and expressiveness in order to address the additional requirements of multimedia resources [16]. The project will advance the state of the art by applying ontology-based discourse structure and analysis to multimedia resources.

The remainder of the paper is organized as follows: in section 2, an overview of the aceMedia project is given. Section 3 describes in more detail the knowledge-assisted analysis, while in section 4, application scenarios using the described framework are described. Finally, conclusions are drawn in section 5.

2. ACEMEDIA PROJECT OVERVIEW

The driving vision behind aceMedia is the opportunity to use knowledge to add value at every step in the multimedia content value chain. A simplified content value chain is shown in Figure 1. aceMedia intends to enable the benefits

This work was supported by the European Commission under contract FP6-001765 aceMedia (URL: http://www.acemedia.org).

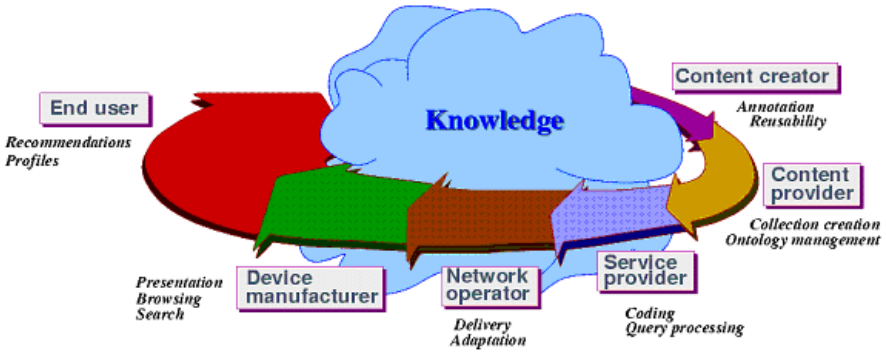

Fig. 1. Simplified aceMedia value chain.

of knowledge extraction and usage to be realised by all the actors in the chain, from content creator to end consumer.

The central concept of the aceMedia Integrated Project is the idea of an Autonomous Content Entity (ACE). An ACE has three layers: the content itself, its associated metadata, and an ”intelligence” layer. The latter consists of distributed functions that enable the content to instantiate itself according to its context including its network environment, the user terminal, and recorded user preferences. An ACE can negotiate and autonomously exchange information with other ACEs, and can update its metadata according to its new status e.g. to show that it has been displayed on a certain user terminal.

The high-level semantic analysis and reasoning framework presented in the following section will help to automatically annotate and generate content descriptors at the ACE metadata layer, and will develop and prototype tools for processing of ACEs such that the derived and associated knowledge can be exploited by the end user in a human-friendly fashion.

In order to achieve these goals, the overall project approach is to divide the work into technical workpackages, which reflect the individual parts of the value chain, and which are linked by the common tasks of system specification and integration:

- **WP2 User:** WP2 will identify user requirements, and translate these into system requirements; apply user models to enable personalisation of the user experience; devise and implement aceMedia user interfaces; evaluate aceMedia with target users.

- **WP3 Content:** WP3 will research advanced content processing to support knowledge extraction; develop scalable coding tools to support self-adaptability of ACEs; research techniques for cross-media adaptation; and implement algorithms to produce enhanced visualisation for navigation and rendering.

- **WP4 Knowledge:** WP4 will enhance the state of the art in knowledge representation to support multimedia content analysis and interpretation; bridge the semantic gap between low-level processing and high-level reasoning; create the ontology infrastructure and tools required for knowledge based audio-video analysis, ontological based text analysis, semantic reasoning, and context analysis.

- **WP5 System:** WP5 will specify the system and architectural framework of aceMedia, including the ACE and Application Module (AM) specifications, and will implement the networking and database components.

- **WP6 Integration and Applications:** WP6 will develop the applications for user evaluation. Two research prototypes will be developed: Personal Content Services (PCS) and Commercial Content Management (CCM).

These technical workpackages are supported by appropriate management, exploitation and dissemination activities.

3. ONTOLOGY INFRASTRUCTURE, KNOWLEDGE-ASSISTED CONTENT ANALYSIS, SEMANTIC REASONING AND INTELLIGENT CONTENT RETRIEVAL

From a knowledge point of view, research has been carried out for several decades in the domains of knowledge representation and reasoning with knowledge. One of the goals of explicit knowledge representation techniques is to provide the ability to derive new knowledge by using logic and inference rules. Over the last decade, this research area has gained new interest in the context of the Semantic Web. New languages such as RDF (Resource Description Framework) and OWL (Web Ontology Language) are currently defined by the World Wide Web consortium (W3C) in order to add meaning to information on the web and allow for

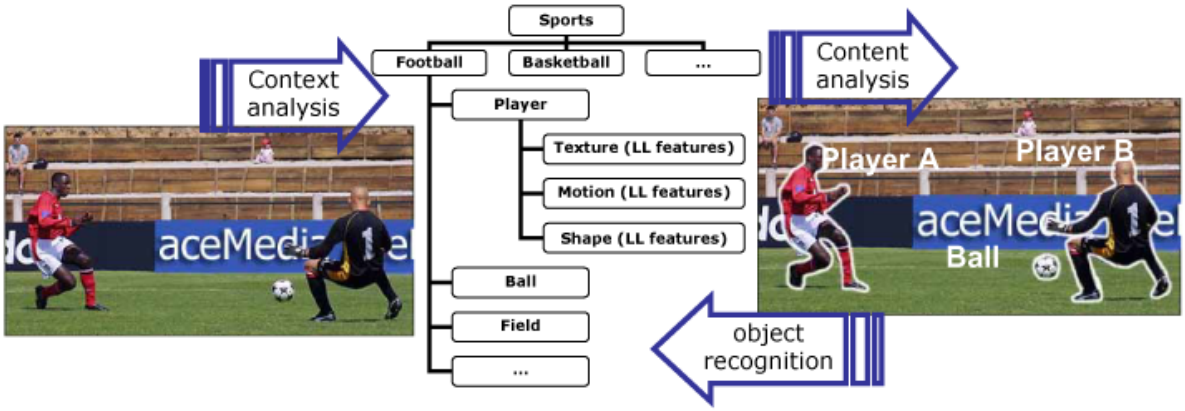

Fig. 2. Knowledge and context-assisted content analysis and reasoning based on a multimedia ontology infrastructure.

better search and retrieval. As a next step, inference rules and logic are to be used by intelligent applications to derive new information from existing information on the web. Ontologies define a set of meanings for a specific domain of information.

aceMedia will investigate ontology modelling in terms of both methodology and expressiveness in order to address the additional requirements of multimedia resources. The project will advance the state of the art by applying ontology-based discourse representation and analysis based on semantic reasoning to multimedia resources. Specifically, aceMedia will:

1. study and evaluate state of the art knowledge representation technologies

2. provide an ontology infrastructure for automated knowledge extraction and usage in multimedia content creation, management, exchange and consumption, and

3. build customized ontologies for the specific domains of the aceMedia applications.

Ontologies will be extended and enriched to include low-level audiovisual features, descriptors and behavioural models in order to support knowledge-aware multimedia analysis, media-aware semantic inferencing and high-level semantic reasoning, which will provide the means for automatic content annotation and generation of the ACE scalable metadata layer. Finally, innovative techniques for context modelling and analysis will be developed, assisting both the ACE creation process and its usage through user query interpretation and intelligent search, retrieval and relevance feedback. The definition, representation and analysis of context will be based on formal concept analysis (FCA) theory, in combination with neural networks and fuzzy set theory.

4. APPLICATION SCENARIOS

In this section, we present an application which illustrates use of the ACE concept, and its value in heterogeneous communication environments.

A media company's marketing director creates a game in partnership with internet service providers, mobile network operators, and interactive TV companies. In the game, subscribers have to watch a short piece of multimedia content, or view a series of still images captured from footage of recent sporting events, or listen to a short audio commentary from one of the events, and decide what happened next. Each game lasts only one day, with the answer and new content available from midnight. Over a period of three weeks, players' scores are accumulated, and prizes are offered to the highest scores.

Manual detection of interesting events to be used in the game (e.g. a kick resulting in a goal during a football match) is a demanding and resource intensive task. The ACE concept provides for automatic and semi-automatic semantic understanding of sports video. It provides shot classification into predefined categories, each of which has a clear semantic meaning (e.g. shots on goal in football video, peleton crash in cycling etc). The ACEs representing these events are self-categorized in relevant catalogues or are selected from automatic matching with user profiles.

In particular, ontology driven multimedia analysis, as illustrated in Figure 2, enables highly efficient processing of stored multimedia material, with specific domain ontologies being applied to the different sports which would be used in the game. Hierarchical event ontologies [17] are especially useful in directing analysis to find the build-up to events of interest, such that the final key event can be omitted for the purposes of the game. The ACE concept also enables unannotated archive material to be used in the game. Preselection of suitable content can take place in an efficient

way by extracting from already used ACEs a machine-level description of the desired types of content (i.e. what constitutes "interesting" in this context), and using it to direct the analysis of videos from the archive.

The aceMedia system would then stream semantically annotated, unfinished, interesting events to potential game players. For this purpose, the game requires content which can be encoded just once, but distributed and viewed in many different formats on many different devices. ACEs support this requirement by using scalable encoding, which leads to the ability of the content to render itself in the appropriate format for the chosen target device. Various ACEs are created such that self-advertising versions of the game, with a previously used piece of content, can be pushed to different terminals (e.g. set-top box, mobile phone, home PC etc).

5. CONCLUSIONS

In this paper an overview of the aceMedia Integrated Project was presented with a focus on the objective of knowledge-assisted analysis of multimedia content. aceMedia targets knowledge discovery and embedded self-adaptability to enable content to be self organising, self annotating, self associating; more readily searched (faster, more relevant results); and adaptable to user requirements (self reformatting). aceMedia introduces the novel concept of the Autonomous Content Entity (ACE), which has three layers: content, its associated metadata, and an intelligence layer consisting of distributed functions that enable the content to instantiate itself according to its context (e.g. network, user terminal, user preferences). In aceMedia, ontologies will be extended and enriched to include low-level audiovisual features, descriptors and behavioural models in order to support automatic content annotation. The ACE concept was illustrated by an example of use in both professional and consumer domains.

6. REFERENCES

[1] S.-F. Chang. The holy grail of content-based media analysis. *IEEE Multimedia*, 9(2):6–10, Apr.-Jun. 2002.

[2] T. Sikora. The MPEG-7 Visual standard for content description - an overview. *IEEE Trans. on Circuits and Systems for Video Technology, special issue on MPEG-7*, 11(6):696–702, June 2001.

[3] A. Yoshitaka and T. Ichikawa. A survey on content-based retrieval for multimedia databases. *IEEE Transactions on Knowledge and Data Engineering*, 11(1):81–93, Jan/Feb 1999.

[4] W. Al-Khatib, Y.F. Day, A. Ghafoor, and P.B. Berra. Semantic modeling and knowledge representation in multimedia databases. *IEEE Transactions on Knowledge and Data Engineering*, 11(1):64–80, Jan/Feb 1999.

[5] A. Yoshitaka, S. Kishida, M. Hirakawa, and T. Ichikawa. Knowledge-assisted content based retrieval for multimedia databases. *IEEE Multimedia*, 1(4):12–21, Winter 1994.

[6] V. Mezaris, I. Kompatsiaris, and M.G. Strintzis. An Ontology Approach to Object-based Image Retrieval. In *Proc. IEEE Int. Conf. on Image Processing (ICIP03)*, Barcelona, Spain, Sept. 2003.

[7] G. Tsechpenakis, G. Akrivas, G. Andreou, G. Stamou, and S.D. Kollias. Knowledge-Assisted Video Analysis and Object Detection. In *Proc. European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems (Eunite02)*, Algarve, Portugal, September 2002.

[8] M. Wallace, G. Akrivas, P. Mylonas, Y. Avrithis, and S. Kollias. Using Context and Fuzzy Relations to Interpret Multimedia Content. In *Proc. 3rd International Workshop on Content-Based Multimedia Indexing, CBMI'03*, Rennes, France, September 22-24 2003.

[9] M. Ramesh Naphade, I.V. Kozintsev, and T.S. Huang. A factor graph framework for semantic video indexing. *IEEE Trans. on Circuits and Systems for Video Technology*, 12(1):40–52, Jan. 2002.

[10] W. Chen and S.-F. Chang. VISMap: an interactive image/video retrieval system using visualization and concept maps. In *Proc. IEEE Int. Conf. on Image Processing*, volume 3, pages 588–591, 2001.

[11] S.S.M. Chan, L. Qing, Y. Wu, and Y. Zhuang. Accommodating hybrid retrieval in a comprehensive video database management system. *IEEE Trans. on Multimedia*, 4(2):146–159, June 2002.

[12] O. Mich R. Brunelli and C.M. Modena. A survey on video indexing. *Journal of Visual Communications and Image Representation*, 10:78–112, 1999.

[13] P. Wittenburg D. Thierry and H. Cunningham. The Automatic Generation of Formal Annotations in a Multimedia Indexing and Searching Environment. In *Proc. ACL/EACL Workshop on Human Language Technology and Knowledge Management*, Toulouse, France, 2001.

[14] H.-P. Schnurr M. Erdmann, A. Maedche and S. Staab. From Manual to Semi-automatic Semantic Annotation: About Ontology-based Text Annotation Tools. In P. Buitelaar and K. Hasida, editors, *Proc. COLING 2000 Workshop on Semantic Annotation and Intelligent Content*, Luxembourg, August 5-6 2000.

[15] J. Wielemaker A.Th. Schreiber, B. Dubbeldam and B.J. Wielinga. Ontology-based photo annotation. *IEEE Intelligent Systems*, May/June 2001.

[16] J. Hunter. Adding Multimedia to the Semantic Web: Building an MPEG-7 Ontology. In *Proc. The First Semantic Web Working Symposium, SWWS'01*, Stanford University, California, USA, July 2001.

[17] C. Dolbear and M. Brady. Soccer highlights generation using a-priori semantic knowledge. In *Proc. IEE conference on Visual Information Engineering, VIE2003*, University of Surrey, UK, July 2003.