An Approach to Self-Annotating Content

Adrian Matellanes, Freddy Snijder and Barbara Schmidt-Belz

Abstract—**aceMedia content analysis capabilities are centered around the concept of the ACE. The ACE is composed of a content layer, a metadata layer and an intelligence layer. In this paper we show one application of the ACE Intelligence layer and how its proactive conduct can help in the complex task of adding semantic metadata to multimedia content.**

Index Terms—**multimedia content analysis, proactive content, self-annotation**

I. INTRODUCTION

Digital multimedia content management is a very complex task. Huge efforts in research and development are being carried out in industry and academia to alleviate this complexity and bring solutions to help end-users and professionals to easily manage their collections of multimedia content.

The aceMedia project tries to help tackle this problem with a wide range of technologies, from multimedia knowledge representation [6], multimedia content analysis [3], personalized search and browsing [4] to content adaptation [8] to cite a few references. Fundamental to aceMedia's approach is the introduction of the Autonomous Content Entity (ACE) [1]. The ACE is a multimedia object comprising three layers : the first layer is the multimedia content itself, the second layer is the metadata layer, which includes manual and automatic annotations, and the third layer is a programmable layer called "Intelligence layer" that provides proactiveness to the ACE.

The intelligence layer is envisaged to help in the complex problem of multimedia content management by enabling the content items to perform actions on behalf of the user, wherever they reside.

This paper briefly describes one of the applications of content proactiveness enabled by aceMedia technologies, namely, the creation of *self-annotating content*.

Content autonomy is not limited to self-annotation, other activities based on the ACE intelligence layer are also carried out [9].

II. MOTIVATION

A. Content flow

Nowadays, content flows from one device to the other. Devices are connected to Personal Area Networks (PANs), Local Area Networks (LANs) and Wide Area Networks (WANs) allowing multimedia content to be easily transferred and shared. Devices supporting multimedia content range from powerful media servers to desktop PCs, set-top-boxes (thin and thick) and small devices such as mobile phones and PDAs. These devices have different characteristics, specific uses and processing capabilities.

This fact about content flowing from one device to the other motivates the idea of a content item (an ACE in the aceMedia context) self-annotating itself whenever it reaches a target device, given that that device has annotation capabilities.

There are different scenarios where content can enrich its metadata as it moves from one device to the other. Different annotation capabilities can be found in different devices, e.g. device A does not have a certain content classification module that is present in device B, but also the same annotation modules can have different capabilities depending on the device on which they reside, e.g. device A's face recognition module may know a different set of persons than device B.

B. User participation

Purely automatic annotations have a long way to go to provide the user with accurate semantic annotations. The semantic metadata associated with the content can be improved with the help of the user. Some of our studies, contrary to some common beliefs, showed that users are willing to "help the system" with their manual annotations [2].

To incorporate users' manual annotations we will create proactive content that analyzes its own automatic semantic annotations and asks the user pertinent questions to solve some ambiguities or add some unknown information, e.g. a face that has been detected is not known to the face recognition module and the ACE asks the user "Who is the person whose face is inside the bounding box?" The user, always in control, can obviously ignore these questions as the system is not strictly depending on them.

III. THE SELF-ANNOTATING PROCESS

A. Proactiveness

In the previous section we have seen some motivations for giving autonomy and proactiveness to multimedia content when we try to add manual or automatic semantic annotations to an ACE. It is important to emphasize here that the whole

process of self-annotation and the ultimate decision to add semantic annotation to an ACE resides in the ACE itself.

We will not go into detail about the software architecture that enables ACEs to run their programmable Intelligence layer in order to give them autonomy; a brief description of this can be found in [1].

B. The AnnotationManager runs the analysis

Content is analyzed by different content analysis modules that produce semantic metadata which in turn is added to the ACE metadata layer. A typical, application driven, annotation process is described in [2].

In our case of self-annotation, it is important to clarify that the ACE intelligence layer is in charge of starting/stopping the annotation process and decides which, if any, content analysis needs to be run. The ACE programmable intelligence layer does not include the analysis algorithms that analyze and produce new metadata.

As explained in the previous section, the modules in charge of analyzing the content and adding new metadata can differ from one device to the other and are offered to the ACE intelligence layer through a common framework called the Annotation-Manager.

This AnnotationManager interacts with the ACE intelligence layer and runs the requested analysis modules in the appropriate order. The AnnotationManager also deals with dependencies, e.g. a face recognition module may depend on a face detection module.

Once the AnnotationManager has called the analysis modules requested by the ACE, it will always run the Multimedia Reasoning module to ensure metadata consistency, remove ambiguities and derive new annotations if possible.

C. Self-annotation process

In this section we will describe a typical self-annotation process. As explained in the previous section, the Self-Annotating ACE is in control of the annotation process but it does not perform the analysis nor the annotations itself. This way, the ACE can benefit from the different capabilities offered by different devices and contexts, see section II.

When an ACE is transferred to a different device, its self-annotation process is started.

The self-annotating Intelligence layer, asks the device what semantic annotation capabilities are present in the device. This request is received by the AnnotationManager.

The AnnotationManager will analyze the kind of content stored in the ACE, i.e. whether it is a still image, a video clip, or any other type of media. Based on this analysis the AnnotationManager decides what analysis capabilities it can offer, e.g. face detection, face recognition, speech recognition, knowledge-assisted analysis, etc.

The ACE checks if this type of annotation has already been performed, and creates a list of the missing annotation categories it is interested in. The ACE sends this list to the AnnotationManager that calls, in the appropriate order, the analysis modules and solves dependencies if needed. The resulting semantic annotations are then added to the ACE metadata layer and finally, as explained before, the Multimedia Reasoning module is called.

We have just outlined a very simple implementation of a self-annotation ACE. The ACE self-annotating intelligence layer can indeed be programmed to perform more complex tasks and take other decisions such as raising questions to the user (see Section II) or prevent certain analysis to be performed (because of privacy issues for example).

IV. CONCLUSION

One of the objectives of aceMedia is to explore advanced content management techniques through the concept of the ACE and its Intelligence layer. aceMedia has successfully created a framework for the deployment of Autonomous Content Entities (ACEs). These ACEs can have proactive behavior that helps users in their digital media management. We have briefly outlined in this paper one of the applications of the ACE Intelligence layer, the creation of self-annotating content. We have presented the motivation which led us to make ACEs self-annotating as opposed to being annotated passively. Finally we have outlined the process and workflow of self-annotation.

Within aceMedia we are investigating other applications of the ACE Intelligence layer that are outside the scope of this paper such as self-organizing ACEs and self-governing ACEs.

REFERENCES

[1] A.Matellanes, T.May, F.Snijder, P.Villegas, E.O.Dijk and A.Kobzhev, "An architecture for multimedia content management" in *EWIMT 2005, London, UK.*

[2] A.Matellanes, A.Evans, B.Erdal "Creating an application for automatic annotation of images and video" in *SWAMM 2006, Edinburgh, UK.*

[3] N.Dalal and B.Triggs "Histograms of Oriented Gradients for Human Detection" in *IEEE CVPR 2005, San Diego, USA, June 2005*

[4] S.Bloehdorn, K.Petridis, C.Saathoff, N.Simou, V.Tzouvaras, Y.Avrithis, S.Handschuh, Y.Kompatsiaris, S.Staab and M.G.Strintzis, "Semantic Annotation of Images and Videos for Multimedia Analysis" in *ESWC 2005, Heraklion, Greece, May 2005*

[5] D. Vallet, M. Fernández, P. Castells, P. Mylonas and Y. Avrithis, "Personalized Information Retrieval in Context" in *MRC 2006 at AAAI 2006, Boston, USA, July 2006*

[6] K. Petridis, S. Bloehdorn, C. Saathoff, N. Simou, S. Dasiopoulou, V. Tzouvaras, S. Handschuh, Y. Avrithis, I. Kompatsiaris and S. Staab, "Knowledge Representation and Semantic Annotation of Multimedia Content" in *IEE Proceedings on Vision Image and Signal Processing, Special issue on Knowledge-Based Digital Media Processing, Vol. 153, No. 3, pp. 255-262, June 2006.*

[7] J.Malobabic, H.Le Borgne, N.Murphy, N.O'Connor, "Detecting The Presence Of Large Buildings" in *Natural Images 4th International Workshop on Content-Based Multimedia Indexing, CBMI 2005, Riga, Latvia, 21-23 June 2005*

[8] N. Sprljan, M. Mrak, G. C. K. Abhayaratne, E. Izquierdo, "A Scalable Coding Framework for Efficient Video Adaptation" in *Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2005), Montreux, Switzerland, April 13-15, 2005*

[9] P.Charlton and J.Teh, "A self-governance approach to supporting privacy preference-based content sharing in distributed environments" in *SOAS 2006, Erfurt, Germany, 18-20 September 2006.*