Offering Access to Personalized Interactive Video

Giorgos Andreou, Phivos Mylonas, Manolis Wallace and Stefanos Kollias
Image, Video and Multimedia Systems Laboratory
National Technical University of Athens
9, Iroon Polytechneioy, Zographoy, Athens, Greece
e-mail: geand@image.ntua.gr

***Abstract*—Semantic treatment of multimedia content, as well as of user preferences and interests, forms one of the most important topics in modern audiovisual content providing systems and opens a great variety of discussions among the research community. In this framework, extracting and combining profiling information about users and content is a challenging problem. This paper is part of our ongoing work in the field of content analysis and user profile extraction towards an interactive video environment; it extends on previous works on low level multimedia content, in the direction of automated extraction of semantic user preferences. Such preferences are utilized towards personalization of the multimedia content offered to the end-users. The latter issue is tackled within the scope of this paper, utilizing a novel mechanism of defining weights that are crucial in the process of automatically extracting profiling information.**

***Index Terms*— personalized multimedia system, profiles, interactive video**

I. INTRODUCTION

One of the main challenges, on which current research area in the field of multimedia content profiling is constructed, is without any doubt the fact that multimedia content retrieval is by far more difficult to tackle than retrieval within the scope of plain text documents; it is much more difficult to extract semantic information from the former. Thus, matching user requests to available multimedia documents is a more complex task to confront. This remark stresses the role of user profiles as extremely important [7].

Three major areas absorb the focus of technological attempts in the field of combining user interests with audiovisual archives and multimedia documents: (i) analysis of a multimedia document for the extraction of specific relevant features, (ii) mapping of these features to user preferences and (iii) presentation of suitable content

This work was partially funded by the IST-2001-34755 MELISA project. This work has been partially funded by the HERACLETUS project of the Greek Ministry of Education.

to the end users, according to their personalized preferences.

Analyzing the content of a multimedia document is a task dominated by high complexity issues. First of all, due to the fact that the entities that lie within multimedia documents and need to be indexed are not directly encountered in the document, recognizable features must be extracted and matched to respective ones found in a specific knowledge base. Additionally, a multimedia document contains objects and describes events, whose relations are spatiotemporal instead of solely grammatical ones. Subsequently, suitable notions must be introduced for the final "balanced" profiling information to be utilized, towards extraction and evaluation of user profiles, based on usage history patterns and preferences. The latter obviously requires intelligent semantic interpretation of the multimedia content.

The structure of this paper is as follows: in Section 2 the foundations of our approach upon previous and ongoing research work performed in the fields of object detection [3] and multimedia content analysis [4] are presented. Based on these, in Section 3 we build upon user profiles, according to information extracted from their semantic preferences and usage history patterns and perform an initial user categorization. The latter is furthermore extended by introducing specific weights applied to system users; this concept, combined with the extracted profiling information, forms an integrated TV entertainment experience for the end users. Finally, in section 4 we present the final content adaptation procedure and a notion of the implementation interface of the so far described system, whereas in section 5 we provide our concluding remarks and summarization.

II. MULTIMEDIA CONTENT ANALYSIS

The main issue that complicates the process when analyzing multimedia documents arises from the existence of objects and events with spatiotemporal relations within them. Abstract concepts like "sports" are not directly encountered in multimedia documents and thus must be inferred from the concrete objects and events and features (like lighting) which are not attributed to a particular object or event in the content.

Several algorithms have been implemented for detecting semantic information using the raw or encoded signal of video material [3]. Although traditional techniques such as scene/shot detection, moving object detection, object feature extraction already contributed a lot in the field of computer vision, semantic-based object detection techniques form the most recent approaches and the forthcoming challenge. Following the task of extracting the bounding polygons and/or the contours of the detected moving objects in a video segment, visual descriptors are utilized to characterize the captured objects or regions, according to the MPEG-7 framework [9].

Subsequently, matching of these descriptors to ones stored in a semantic knowledge base is used as

a means for automatic detection of events and objects. In the following a representative set of these descriptors are briefly presented and categorized, as follows:

- **Color (RGB, HSV, Grayscale)**: Variance, Percentage/Histogram, Spatial Coherency (Structure), Spatial Distribution.
- **Shape**: Region Shape by ART (Angular Radial Transform) Coefficients, Curvature.
- **Motion**: Motion Activity, Intensity of Activity, Direction of Activity, Spatial Distribution of Activity, Temporal Distribution of Activity, Spatial Localization of Activity, Rigidity.
- **Localization**: Region Locator, Spatio-temporal Locator.
- **Edges**: Edge Histogram.

A "semantic index" is generated as a result of automatically recognizing objects and mapping them to "semantic entities" for a multimedia document [2]. This issue is in fact very complicated and still open; similar input, though, can be acquired through textual analysis of the structured textual information contained in the metadata information that accompanies the annotated multimedia material. In this work we use a semi-automatic index generation mechanism for multimedia documents, as well as automatic index generation algorithms for textual documents, as the primary step towards the multimedia content analysis.

III. Extraction of User profiles

Current personalization scenarios accompany several interactive services, applications and content delivery to the end-user [1], [5], [8]. In this work, we focus on "content adaptation" through the underlying creation of the user profiles, where different multimedia content can be generated for individual users or playback devices, such as mobile phones, mobile computers, PDAs, etc.

Several issues need to be taken into consideration, while constructing the initial profiles. The profile building process is mainly based on specific user interests, the offering of audio and visual enhancements, i.e. offline and in-play statistics, distance and position indicators, interactivity notions, i.e. online betting and advertising and several event alerting functionalities, related to the former information (e.g. a goal scored in a football match, while the user is viewing an athletics broadcast). Moreover, the main points to consider, regarding the specific user attitudes, may be summarized in the following:

- A user may be interested in multiple topics.
- Not all topics that are related to a multimedia document in the usage history are necessarily of interest to the user.

All of the above issues are tackled using similar tools and principles, as the ones used to tackle corresponding problems in content analysis [10]. Thus, once more, the basis on which the extraction of preferences is built is the context [4]. The

common topics of multimedia documents are used in order to determine which of them are of interest to the user and which exist in the usage history coincidentally. Since a user may have multiple interests, we should not expect all documents of the usage history to be related to the same topics. Quite the contrary, similarly to semantic entities that index a document, we should expect most documents to be related to just one of the user's preferences.

Therefore, a ranking of documents, based on their common topics, needs to be applied. In this process, documents that are misleading (e.g. documents that the user chose to view once, just to find out that they do not contain anything of interest to him) will probably not be found similar with other documents in the usage history, resulting to only those documents that form the desirable distinct user profile categorization.

Finally, in order to balance the obtained results from the above user profiling procedure, we introduce the concept of "user weights". These weights are predefined within the system specifications and categorize its users into three main categories: *headstrong users*, *intermediate users* and *unconcerned users*. The combination of the profiling information and the stereotyped weights, results into a descriptive and representative profile for each user group in a "Least Misery Strategy" [6]. This aggregate profile will then be utilized to provide them with content of their interest, adapted to their liking.

IV. Content Adaptation & Personalized Interface

In this section, the final content adaptation procedure is presented, as well as a representative chunk of the system implementation. At this step, all previously extracted information from the usage profiling process is combined with the visual enhancements and metadata information. This adaptation procedure must be performed because of the diverse nature of the client devices and the specific kinds of the enhancements themselves. It is based on the content analysis already performed, as well as on the users' usage history manipulation. As a result, the final audiovisual content presentation to the end user evolves continuously according to the supplied profile information.

In the EU IST project MELISA, where he methodologies presented in this paper are followed, separate initial presentations are built for separate user classes, according to their extracted user profiles. Figure 1 shows the initial presentation for an intermediate user profile terminal, regarding a soccer game scenario.

Figure 1. Initial intermediate profile soccer game scene

One of the most important interactive functionalities common to all scenarios targeting end users belonging to this intermediate profile is the Bet Menu, which allows users to place a bet on a player during the particular event. When the event starts, the client platforms receive the full initial scene together with the first images of the event. During the event, the initial scene is enhanced with generated online content with real-time data. This allows the program director to modify the parameters of the scene at any time, according to what happens in the event. A typical example of real-time content change is the modification of the Bet Menu. Typically the odds for an available bet would vary during the event. Figure 2 shows the result of several generated real-time contents on the initial presentation given in Figure 1.

Figure 2. Updated intermediate profile soccer game scene

As we see on the top right corner of Figure 2, the score has been changed, as well as the current game time. Also, following an important game action, a new image object has been displayed on top of player 9 which, when selected (i.e. by pressing 9 on the remote control) shows statistics for this player. Practically, this corresponds either to the specific user actions per system terminal, or to the specific profile class descriptions.

Towards the interface resulting from such a combined adaptation, regarding profiling and extra available information, we present in particular two different screenshots from a specific MELISA software client, indicating the differences between the user interfaces, as well as the ones between the content enhancements arriving at the end-user.

Figure 3. Plain client

Figure 4. Enhanced client

In Figure 3 we observe a plain user interface with minimal extra information, whereas in Figure 4 several visual enhancements and interface changes are present (betting and replay buttons, as well as player number indicator), providing more content management powers to the end–users and indicating the existence of a more enhanced user profile. In this scenario, users press the relevant colored button on their remote to access these features or view more information on the tracked player.

V. Conclusion

This paper builds on previous work on low- and higher-level multimedia document analysis by utilizing user preferences and user weights towards an overall personalized multimedia viewing experience, concerning the accompanying metadata

information and end-user interface. Such preferences are utilized together with applicable user weights towards the personalization of the multimedia retrieval process. The methodology, techniques and system presented herein have been developed in the framework of the EU IST MELISA project.

REFERENCES

[1] Correia N., Boavida M., Towards an Integrated Personalization Framework: A Taxonomy and Work Proposals, Workshop on Personalization Techniques in E-Publishing on the Web, Spain, 2002.

[2] Zhao, R. and Grosky W.I., Narrowing the Semantic Gap-Improved Text-Based Web Document Retrieval Using Visual Features, IEEE Trans. on Multimedia, Vol. 4, No 2, June 2002.

[3] Tsechpenakis G., Akrivas G., Andreou G., Stamou G. and Kollias S., Knowledge - Assisted Video Analysis and Object Detection, Symposium on Intelligent Technologies and their implementation on Smart Adaptive Systems, Albufeira, Portugal, September 2002.

[4] Wallace M., Mylonas P., Kollias S., Automatic Extraction of Semantic Preferences from Multimedia Documents, International Workshop on Image Analysis for Multimedia Interactive Services, Portugal, 2004.

[5] Spangler W., Gal-Or M., Masy J., Using data mining to profile TV viewers, Communications of the ACM, Vol. 46, Iss. 12, pp. 66 – 72.

[6] Masthoff J., Group Modeling: Selecting a Sequence of Television Items to Suit a Group of Viewers, User Modeling and User-Adapted Interaction, vol.14, no.1, pp.37- 85.

[7] Angelides, M.C., Special issue on Multimedia content modeling and personalization, IEEE Multimedia 10(4)

[8] Wallace M., Karpouzis K., Stamou G. Moschovitis G. Kollias S., Schizas C., The Electronic Road: Personalised Content Browsing, IEEE Multimedia 10(4), pp. 49-59, 2003 [With correction in 11(1), pp. 53-53, 2004]

[9] Sikora T., “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.

[10] Akrivas G., Wallace M., Andreou G., Stamou G. and Kollias S., Context – Sensitive Semantic Query Expansion, ICAIS, Divnomorskoe, Russia, 2002.