Handling multiple channel video data for personalized multimedia services:

a case study on soccer games viewing

C. Patrikakis, N. Papaoulakis, C. Stefanoudaki, A. Voulodimos, E. Sardis
School of Electrical and Computer Engineering
National Technical University of Athens, Greece
{bpatr, npapaoul, cstef}@telecom.ntua.gr, {thanosv, sardism}@mail.ntua.gr

Abstract— **Personalization is undoubtedly present in today's pervasive and ubiquitous environments and tends to be an increasingly popular requirement in every technological aspect of the everyday life. When it comes to live multimedia services, however, there is still a lot to be done to satisfy the demanding user who wishes to view the most important parts of events that take place in parallel. In this paper, a service that allows users to personalize the provision of multimedia streaming services offered through broadcasting networks is described. The service is aimed at offering full personalization capabilities over traditionally broadcast oriented terminals such as TV sets and Set Top Boxes. Complete personalization can be achieved through analog TV or DVB broadcasts with the use of personal devices such as mobile phones by using IP connection for the return channel (RC) (DVBRC). A use case scenario featuring one example related to personalized viewing of two parallel athletic events (such as Champions League soccer matches) is described and used for the evaluation of a prototype implementation of the service.**

Keywords- Multimedia annotation, personalization, metadata, broadcasting, multiple channel video data.

I. INTRODUCTION

Broadcasting is fast declining. Nevertheless, new technologies using this type of transmission are coming to the fore with TV widgets being the most recent example.

The rapidly increasing capabilities of networking infrastructures and end user devices for supporting high bitrate multimedia streaming services has led to a corresponding increase in the demand for real time multimedia services such as Web TV and Internet TV [1]. A plethora of devices ranging from TV like Set-Top-Boxes up to mobile phones constitutes a wide set of platforms used for access to high quality content transmitted over a corresponding wide range of wired and wireless access networking technologies. This has led to the introduction of added value multimedia services, offering enhanced user involvement (i.e. fusion of user generated content) and context awareness. The need to offer end users quality aware multimedia streaming that can be adapted to network conditions so as to support enhanced Quality of Experience, together with the ability to support sophisticated mechanisms for user preferences capturing, have led to the development of the tools upon which enhanced user control over the offered services can be based. Seamless bitrate switching, transparent transitions between access network technologies during media streaming and automatic, context aware adaptation of transmission parameters are characteristic examples as regards the networking part. As regards the management of the transmitted content, the existence of advanced image - video annotation [2] and characterization mechanisms, allows the automatic (even real time) [3] annotation of multimedia content so that truly personalized media streaming can be offered based on the matching of user selections to content characterization [4].

An important use case scenario, for the provision of personalized, highly interactive multimedia services of particular interest is athletic events coverage, such as the Olympic Games [5], [1]. In this case, allowing users to have full access on and control to the multimedia content transmitted to this terminal signifies a transition from traditional TV-director managed broadcasting to a user and interactive based management of the viewing experience.

II. THE NEED FOR PERSONALIZATION

Video consumption across Internet connected devices continues to rise. In fact, the recent developments in Internet and multimedia technologies, have led to the definition of the "three screens" concept, which are deployed on a daily basis to provide access to media broadcasted content: the television, the personal computer and the mobile phone [6]. Here, the role of TV still remains important; according to recent reports, in U.S.A. there are more TVs in each home than people [7]. At the same time, TV viewing habits are definitely changing, with an increase of 40% as regards the use of Personal Video Recorders, while mobile video viewing has witnessed a 50% increase in 2009 compared to the previous year [7]. This has created new scenery, with great interest in the personalized provision of broadcasting services, giving emphasis on personalizing the way that live coverage of events is provided.

In this context, coming to the issue of live sports coverage, it is a common practice in league tournaments or championships that several matches take place in parallel, for several reasons: for minimizing the time needed for all matches to be completed or avoiding expedience play by some teams. Since the matches are broadcasted by the TV stations, there are two options: Either to broadcast the matches live, using several channels (DVB bouquet), or to broadcast one match live, and the others using a recorded version. The first case does not allow the viewers to watch all matches in which they would be interested using their home electronic devices; even the use of Picture in Picture (Picture in Picture - PiP), i.e. the feature that allows one program/channel to be displayed on the full TV screen at the same time as one or more other programs/channels are displayed in inset, smaller windows, offered in some TV sets is not a solution, since it is difficult to focus on the coverage of two matches in parallel, while the feature covers only up to two matches. The PiP feature for example offers audio provision only for the main channel and the analysis of the small screens is very low and can be used only for monitoring. The second option, though it allows viewing of all matches, is supported in real time only for one of them, while in several cases the commentator may have announced a priori all results, thus reducing the interest of the viewers for the recorded matches.

During a large athletics meeting, a big number of broadcast cameras are placed around the stadium in both static and mobile modalities, arranged for the best coverage of the actions. In the conventional broadcast networks, the human director selects the camera sequence and instructs each of the cameramen to capture the activities based on the sports schedule and their personal experience. Only a very small proportion of the captured videos is usually passed on to the viewer at home; the remainder is either temporarily archived or deleted. In IST My-e-Director 2012 FP7 project, video perception modules have been developed to intensively scan in depth all of the raw video feeds, before human editing, in the hope to extract interesting and informative sequences that would otherwise be discarded [8]. The rationale behind the project's platform is simple: If we can automatically localise and identify athletes in video feeds, we can potentially provide users with a variety of interesting features, such as smart zoom (pan and scan), the choice of the best un-occluded view of an athlete, etc. In the rest of this paper, we present the use of My-e-Director 2012 platform which can act as a personal director for each user by following his preferences. Though the platform is capable of supporting switching among heterogeneous access networks and end user devices, in this paper, emphasis is given in a particular scenario: that of using a TV in combination with a personal device (a PDA) in order to personalize broadcasting services offered over DVB [9]. The enhanced capacity of DVB networks offers a perfect broadcast solution for hosting the additional camera streams. The use case scenario selected for the presentation of the platform deployment is that of a European Cup soccer night. For a comprehensive description of the platform, the reader is directed to the web site of the project [8].

III. PLATFORM ARCHITECTURE

In order to understand the functionality of the platform, the following Figure 1 represents the information and video flow from the sports venue, up to the end user's device. It should be noted that some parts of the architecture such as the Broadcaster's infrastructure for managing the video broadcast are presented in a simplified way, since they are of no interest and have no impact on the design.

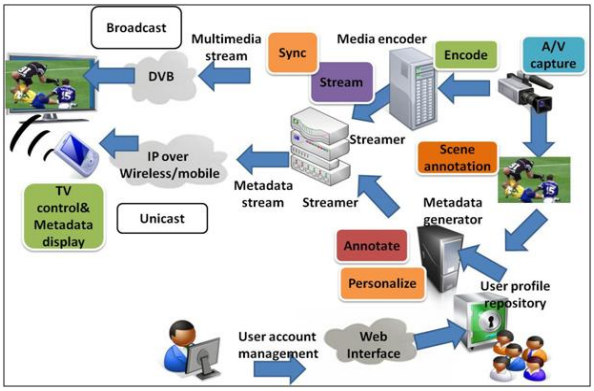

Figure 1. General platform architecture overview

This live video and audio content from the cameras covering the athletic events (in our case the soccer games which take place in parallel) is encoded properly (i.e. MPEG4/H.264 format) in order to be transmitted to the media streamers. In addition, as a parallel or preceding procedure, metadata containing information about the games (such as statistics and events that take place in parallel with the action in the field) are generated. This metadata information is filtered through the selection of the most important events and favorite athletes according to user preferences registered in their profiles. In other words, the metadata are created based on users' favorite events and players. More on the process of metadata creation will be given later in the paper.

To support the streaming procedure, there are two streamers: the metadata streamer that takes the generated aforementioned metadata and encapsulates them in packets so that they will be transmitted to the end user and the multimedia streamer that is used to stream the corresponding video and audio data. The two different streams are

synchronized before being delivered to the end user. After the synchronization, the multimedia streams are broadcasted over DVB to the end users' TV sets, while the metadata streams are sent in unicast mode over wireless and mobile technologies using IP to the end users' personal devices such as PDAs and mobile phones. The personal device has a double-role. Firstly, it is the device which receives all the metadata about what is happening to all the athletic events that interest the user. Therefore the user is able to receive information about all the soccer matches in real time. Secondly, it is the mean through which the user can control what will be displayed on the TV screen; more specifically it makes the switch between TV channels that both contain important events for the user through the use of the appropriate interface (i.e. infrared or Bluetooth), acting as a remote control.

The architecture allows the users to watch the soccer match that interests them most, while in parallel they get information about what is happening in other matches. However, what is more important is that the information delivered to their device is personalized through the use of personal profiles and therefore matches their preferences, while through the use of the remote control capabilities, automatic switching of the TV channel so as to tune in to the match in which the most important event is happening at each time. An example will make this function clearer: Let us imagine that the user is watching a match, where a free kick is about to take place. In parallel, in another match for which the user has expressed interest in his personal profile, a goal attempt has just taken place. His PDA device receives the related information and displays it, while since the event is important it automatically switches to the corresponding channel, unless the user decides to opt out.

The implementation of the application can be divided in three parts. The first part is the web service which was implemented using the Microsoft's WCF platform and was hosted in an IIS server. The second part is the user interface of the web service, in which the user creates his profile and is able to see the metadata and the recommendations, and it was implemented in two versions: one over the Silverlight platform and one using plain HTML and javascript. The reason for the different versions was in order to test the usability in different platforms (a more elaborate but supported by less devices, and a more easy to deploy in most of devices, following a more traditional approach as regards the user interface). The third and last part is the application which is used to support the channel switching. This was implemented in Windows Mobile platform. The communication is accomplished through SOAP using the HTTP protocol. Using this implementation, we have succeeded in letting the user access the service from any operating system, but in order to enjoy the automatic channel switching service he has to use a windows mobile device capable of switching the channels to his TV set.

IV. USE CASE SCENARIO

A use case scenario used for testing and evaluating the platform is described subsequently. The scenario features the parallel viewing of two soccer matches, where the platform is used in order to provide the user with the capability to follow both matches, while he is informed about the most important events, and whenever something important happens in one match, the TV channel is switched so that the user is able to watch it.

A. User login and user's profile creation

During the login procedure, the user enters his username and password in the corresponding boxes and then continues with the registration of his sport event viewing preferences so as to create his profile. This is made possible through the use of a web server acting as the front end to the user preferences repository. Due to the fact that this specific scenario is related to soccer games, users' profiles are going to be formed by marking their favorite events as regards punishment, goals/ efforts on goal and others.

The screenshot in Figure 2 shows that, there are four levels to characterize the importance that each event has for every user; 1) great, 2) so-so, 3) small and 4) not at all. The concept is that during the soccer game and based on the importance the user has given to each event, he will receive the corresponding annotations. Examples of the events are goals, offsides, dropped balls, throw in, etc. The user, after marking his favorite events, has the chance to select his favorite players.

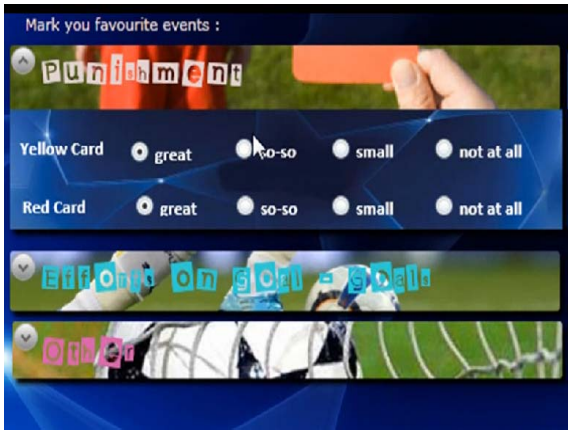

Figure 2. Selection of favorite events

Due to the fact that two soccer games take place at the same time, the user may select among four soccer teams together with the names of their players (please refer to Figure 3). As it is depicted in the following figure, the user ticks the box on the right of the players' names. This action means that he places the chosen players in higher level of importance than the non-chosen ones, so any metadata or recommendations that include them are in priority. By this last action, and the saving of his choices, user completes the login procedure and the profile creation.

Figure 3. Selection of the favorite players

As mentioned above, in order to evaluate the use of the platform, we have simulated a scenario related to two soccer games; G1 and G2. In these games there are two Greek soccer teams (HG1 and HG2) which are competing against other European teams (EG1 and EG2) in the context of an European Cup (i.e. Champions League), with the kick off set at the same time. It should be noted that in order to avoid any problems related to the use of registered trademarks, property rights or broadcasting rights, the names of the teams have been substituted with code names (HG1, HG2, EG1 and EG2), while in the figures used, the logos and names have been blurred. The same applies to the names of players appearing in notification messages. Furthermore, in order to provide the users that have evaluated the service, with look and feel that resembles the graphics and templates used in an actual European Cup, the applications on the mobile device as well as the web server pages for profile selection have been designed accordingly. Finally, in order to simulate the metadata and annotation information in a manner that is as close as possible to the actual match coverage, the way UEFA soccer matches are covered through UEFA web site has been studied [UEFA]. The match descriptions available for free over the internet at the UEFA web site have been used as the basis for the production of the corresponding metadata while for the video, recorded versions of the broadcasts of matches over public TV in Greece were used.

The simulated scenario is the following: A viewer sits in front of his TV set on a European Cup night, having also his PDA nearby. He has selected to watch two matches in parallel; HG1 competing against team EG1 and HG2 competing against team EG2 which play at the same time. The user enters the web site, logs in and selects from the list of available matches, the corresponding two, and also completes his selection by selecting particular players that he wishes to follow. He also checks on his regular preferences as regards the importance of match events (fouls, corner kicks e.a.).

The user starts by watching one of the two matches that he has already selected. During the match, he receives on the PDA screen metadata based on his preferred events and players. The information about the coverage of the matches is filtered so that the produced metadata stream per user matches the user selections made during the profile creation. It should be noted that in order to avoid frequent (thus annoying) notifications, only the important – according to each user's selections - metadata will be sent to the user's personal device

In case an important event happens in match B while the user is watching match A, the user is informed about it with a notification. If the event is so important that a channel switch should take place so that the user can watch what is happening in match B, a corresponding recommendation appears on his PDA screen. After some seconds and if the user takes no action, the PDA makes a channel switch on the TV, so as to tune in to the channel where match B is displayed. Switching between channels may be performed through the use of Bluetooth or Infrared commands sent from the PDA to the TV screen. For the needs of the evaluation, an external application controlling the Infrared port of the PDA was used. The use of Infrared as the interface for TV interfacing was selected in order to evaluate the platform under real life conditions. However, use of BT when this is available in TV sets is easy to implement. The user continues watching the match in the second channel, until a new recommendation (indicating an important event) for switching back in the first channel comes. Alternatively, the user may select to switch manually, using the interface on the PDA, so that the platform is informed about his selection. The procedure repeats until the end of the matches. By the end of the matches, the user has seen the most important events of both soccer games.

The 'background' procedure of the metadata creation is the following: For each sport, we have created a list of events which are possible to happen. For example, for soccer, some of the possible events are the goal, the attempt for goal, the foul near the box, the foul away from the box, etc. Every event is assigned two weights: one according to the user preferences *[event_user_pref]* and another according to how

important this event is *[event_weight]* (e.g. in football the foul away from the box event has a smaller weight than the attempt for goal event). Moreover, we have the user preferences about the athletes, the teams and the sports they like. Every metadata is formulated from the event (from the above lists) that is described, the player who is participating, the team and the sport. We evaluate the importance of the metadata using the following formula:

([event_weight][event_user_pref]+[player_user_pref]* [team_user_pref])*[sport_user_pref]*

B. Identify the Headings

There are many approaches on how the annotation of an important event to the users is created. In the first approach, the creation of annotations can be made manually; this is an easy and very low cost solution. An implementation based on manual annotation has been proposed in the project 'SportScout'. The user-'scout' can enter a specific game in this platform -for example a football or a basketball game- and among other capabilities, to make free commentary on the scenes. SportScout is already used by the football and basketball sections of important Greek teams [10]. Another solution is the use of existing information regarding the coverage of a match, such as the information already available for free at UEFA Champions league site, in which there is streaming textual description of the soccer game being produced in Real Time – Live. The third and state of the art solution is based on real time image and video analysis and tracking, performed through a fully automated procedure of annotation creation. [11]. All the above alternatives result to the creation of a metadata stream that is personalized for each user, and is synchronized with the video before transmission.

V. USE CASE SCENARIO

Students of Electrical and Computer Engineering School of National Technical University of Athens (both undergraduate and postgraduate) were used as testing users to watch two soccer matches, experience the way the platform works and make their own comments about its functionality. Some of the comments led to the conclusion that the platform may become annoying when the user is getting continuous annotations and recommendations of channel switching. Taken into consideration the multiple times recommendations appear, a possible proposal to the users would be that they should lessen the number of different types of events they characterize as important. For example, as regards soccer games which are presented in this paper, users can avoid annoying recommendations if they choose as very important only the event of goal and not both goal and offside for example.

For this, in the future and following a larger scale evaluation study, an extra level of personalization including the frequency of notification dispatching that the user considers logical is under consideration, with some events such as goals and red cards being able to override the selected notification filtering policy.

As regards the issue of extending the platform, an important issue is the integration of logic that allows the system to become self-adaptive to the choices of the user (what he watches indeed, except for what he had chosen in his profile) and his reactions to the recommendations of the system, for example whether he accepts the channel switch that is proposed by the platform. This can result in a platform that is able to make recommendations more dynamic and more adaptive to the user watching behavior; it will 'catch' possible differences between final user watching habits and user choices in the profile.

Another extension worth investigating for usability purposes is the display of recommendations in the form of a video window showing what is happening in the parallel event (i.e. PiP) for TVs that allow it, instead of direct channel change. However, this requires that the PDA device is programmed to use the specific functionality for enabling PiP. Instead of having the form of a countdown clock prior to switching, this alternative would give the user an elaborate view of what is happening and make his decision about accepting the recommendation easier.

ACKNOWLEDGMENT

The research leading to these results has received funding from the European Union's Seventh Framework Programme ([FP7/2007-2013]) under grant agreement n° ICT-215248. The authors thank My-eDirector2012 project participants for their support, which has assisted in the preparation of this paper.

REFERENCES

[1] S. Poslad, A. Pnevmatikakis, M. Nunes, E. G. Ostermann, P. Chippendale, P. Brightwell, Ch. Patrikakis, "Directing your own live and interactive sports channel", presented at 2009 10th Workshop on Image Analysis for Multimedia Interactive Services, London, UK, May 2009.

[2] Dorado, A. ; Izquierdo, E.; Semi-automatic image annotation using frequent keyword mining. In IEEE Proceedings of Seventh International Conference on Information Visualization, 2003, (IV 2003), pp. 532–535, 2003.

[3] Ran Li; Jianjiang Lu; Yafei Zhang; Zining Lu; Weiguang Xu; A Framework of Large-Scale and Real-Time Image Annotation System. IEEE International Joint Conference on Artificial Intelligence, 2009. JCAI '09, pp. 576 – 579, 2009

[4] Jia Li ; Wang, J.Z.; Real-Time Computerized Annotation of Pictures. IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 985 - 1002 , June 2008.

[5] Charalampos Z. Patrikakis, Pantelis Karamolegos, George Koukouvakis, "The Olympic Games: The Way toward the Internet", Annual review of communications, IEC ISBN:1-931695-28-8, Vol. 58, [pp 335 - 340], Nov 2005.

[6] Steve Clayton,"Three screens and a cloud", available at http://blogs.msdn.com/stevecla01/archive /2009/05/21/three-screens-and-a-cloud.aspx, accessed on 22/3/2010.

[7] A2/M2 Three Screen Report, vol.5, 2nd quarter 2009.

[8] My-eDirector 2012 IST project web site, URL: http:// www.myedirector2012.eu, accessed on 12/1/2010.

[9] The Digital Video Broadcasting Project (DVB), "Return Channel Terrestrial", WP-03, 3/2001, available at http://www.dvb.org/technology/fact_sheets/WP03_RCT.pdf, accessed on 28/4/2010.

[10] Nikos Sfingos'SportScout- a few steps guide', Thessaloniki, GREECE, 2009.

[11] My-e-Director 2012-Real-Time Context-Aware and Personalized Media Streaming Environments for Large Scale Broadcasting Applications project, Deliverable D3.3: "Multiple-Person Tracking in an Unconstrained Environment", 2009.

[12] D. Farin, J. Han, P.H.N. de With, "Fast camera calibration for the analysis of sport sequences," ICME, pp. 482-485., 2005 IEEE International Conference on Multimedia and Expo, 2005.

[13] N. Katsarakis and A. Pnevmatikakis, 'Event Detection in Athletics for Personalized Sports Content Delivery', WIAMIS 2009, London, UK, May 2009.

[14] Mario Serafim Nunes, Charalampos Z. Patrikakis and Nikolaos Papaoulakis, "A network oriented perspective on the personalization of media streaming", presented at the 2nd IEEE Workshop on Enabling the Future Service-Oriented Internet of IEEE Globecom 2008 conference, New Orleans, USA, Nov-Dec 2008.

[15] P. Perez, J. Vermaak, A. Blake, "Data Fusion for Visual Tracking with Particles", Proceedings of IEEE, Vol. 92, No. 3, Mar 2004.

[16] A. Pnevmatikakis and L. Polymenakos, 'Subclass Linear Discriminant Analysis for Video-Based Face Recognition', Journal of Visual Communication and Image Representation, Vol. 20, Issue 8, pp. 543-551, Nov 2009.