Collaborative Personalized Digital Interactive TV Basics

Jurij F. Tasič, Janez Zaletelj

University of Ljubljana, Faculty of Electrical Engineering, Ljubljana, Slovenia
Laboratory for Digital Signal, Image and Video Processing
Jurij.tasic@fe.uni-lj.si

Abstract — With the introduction of interactive user oriented TV, digitalization and ipTV, a whole broadcast world is facing new challenges as well as new opportunities. The world of personalized digital interactive TV today is a fast developing area with well established and proven research concepts. Nowadays devices should be designed to learn for themselves and according to the accepted knowledge to prioritize the available programs for the final user. The emergence of the Digital TV technologies calls for investigating consumer preferences and behavioural models towards this new promising communication services. In the paper we are presenting the basic principles of a user based recognition and modelling of personalised interactive digital TV technologies which might increase and create new opportunities and home services.

1. INTRODUCTION

Nowadays TV audience preferences and theirs habits – profiles are mostly ignored by companies offering TV programs. The interactive TV should support the TV viewer in its intentions to be actively involved in the TV content delivery to his TV set, where the intelligence incorporated in the interactive IP TV should identify the user and its habits by interfaces collecting such information. In the case of live, real time transmissions like sport transmission the video conductor has no information who is watching the program and what is happening with the audience, so he is not able to react if the audience is not interested in the content.

The introduction of interactive TV [1,2] thus brings new challenges as well as new opportunities to the existing world of TV services. The interactive TV applications enable the consumer to actively give feedback on the TV program. This feedback can be collected by the application and fed-back to the content delivery office. The information on the audience preferences could be used for the planning of the TV program. However it can be also used during the production of the live program such as sport transmission. This would require that the user feedback is collected in real time and that the whole content delivery process is adapted to effectively use this new information.

Within the 6th Framework Programme project called “Live Staging of Media Events” [3] we are developing together with partners a new concepts of TV production based on iTV technologies. The goal of the project is to provide new workflows and supporting tools for the Video conductor to enable him to react in real-time to live events and to use feedback information from the audience to provide interesting new viewing experiences to the TV audience.

The prosperity of interactive digital TV applications therefore will depend on the new content adaptability and ease of use. The investigation of simple, intuitive human-machine interfaces is therefore crucial for their widespread acceptance. Personalized iTV service basically requires ***consumer authentication*** and consumer ***behavioral models*** with personalized content delivery.

In this paper, we investigate alternative ways based on computer vision to enable consumer authentication for personalized iTV services [5]. The DVB digital broadcasting standard and the MHP platform enable production of multichannel, interactive TV content. Within the mentioned EU-funded project ”LIVE” new concepts and methods are being investigated which will enable on-line - ”live” production and delivery of interactive multichannel content, targeted for the sport domain. This interactive content will be offered to the TV viewer -consumer in a personalized way through a suitable MHP-based user interface. However, to enable delivery of personalized services the consumer must be first authenticated and secondly supported by personalized content delivery.

The normal authentication procedure includes logging in to the set-top box through the remote control, which is not a very user-friendly procedure. In this paper we investigate the use of computer vision methods to automate the process of user authentication and personalized content delivery. In this paper we propose to use a computer vision subsystem on the STB connected to the video camera, which continually monitors and tracks the TV viewers. Figure 1 shows LIVE iTV framework and device setup.

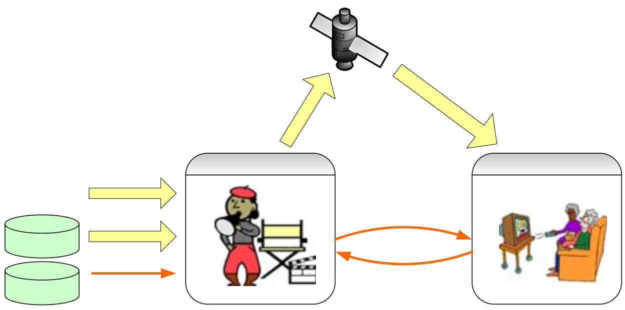

Fig. 1. Device setup.

In the phase of authentication the computer vision subsystem must meet the demanding requirements of the home environment. The main problems of home environment are the following:

- variable lighting in home environment,
- multiple moving persons watching the single TV,
- variable position and gaze of the viewers.

The goal of the image processing subsystem is to continually track and recognize all the people’s faces in front of the TV set, so that the user interface and the TV services could be adapted and personalized in real-time. To achieve this ambitious goal, the proposed system combines multiple image processing algorithms including face detection, eye detection and tracking. Based on the eye detection, face registration and normalization is proposed which increases the recognition rate compared to the standard eigenfacemethod used on non-registered and non-normalized face images.

2. USER AUTHENTICATION USING FACE RECOGNITION

2.1.Face-detection

Face detection such as face recognition are first two crucial steps in the Collaborative Personalized Digital Interactive TV system. The goal of the face detection step is to extract a human face from still images or video and to eliminate disturbing factors such as background, hair etc. from a face image. We implemented a method for multiple face detection developed by Paul Viola and Michael Jones [4], and upgraded by Rainer Lienhart and Jochen Maydt [6].

Method uses Haar-like visual features for face detection and AdaBoost learning algorithm which selects a small number of features from a larger set. Method also combines several weak classifiers in a cascade, which allows uninteresting objects to be quickly discarded while spending more time on promising regions. This method yields high face detection rate at low false rejection rate and runs in real-time. Method was learned to detect faces from frontal side (Figure 2). Face images have to be registered due to differences in size and position in detected faces. This can be done by detecting eyes in a face region and using their coordinates as reference points. We propose simple algorithm for eyes detection, which could also be used for gaze tracking and therefore user feedback collection (e.g. whether user is watching the current TV program or is napping because he is bored).

Fig. 2. Face detection using Haar-like features and AdaBoost algorithm.

2.2. Eye detection for face registration and gaze tracking

Algorithm for eye detection contains two major modules. First module searches for eye candidates by using color information from face images. Second module verifies and selects detected eye candidates by contour analysis. Eyes detection from color images requires selecting an appropriate color space and a cluster associated with a color of eyes in that space. Since this cluster heavily depends on luminance in the RGB color space, we used YCbCr color space, which has luminance component separated fromcolor/chrominance components. An analysis of the chrominance components indicated that high Cb values and low Cr values [7] are found around the eyes (Figure 3). An analysis showed that this area contains low Y values, because iris and pupil are usually darker than corners.

Fig. 3. Y, Cb, Cr components of YCbCr color space.

Eye candidates can therefore be selected by thresholding color components with appropriate threshold and combining obtained binary masks with logical AND operation. Because proper threshold value heavily depends on variable lighting in an environment and also changes from face to face, adaptive thresholding was used. Threshold was set at a value, at which eyes' area spanned around 4% of face area. Morphology-based operation, dilation, was used to remove irregularities, and filtering with median filter was used after combining the masks. Usually more candidates were found by using only color information from image, therefore candidates' contour analysis is also used. Contours were extracted from binary mask. Contour area and center of gravity were calculated for each contour.

Algorithm described was tested on two datasets. The first one is AR dataset [8], which contains face images with white background. Only face images without occlusions and with uniform lighting were used from this dataset. The second is FFD database [9], which contains face images with different real-life backgrounds. Face and eye detection on the images from both dataset is shown on figures 4 and 5. The eye detection rate is presented in table 2. The rate exceeded 90% in both cases with processing time around 23 ms for one image on a computer with Intel Pentium 4 3.2 GHz processor, 2 GB RAM andMicrosoftWindows XP operating system.

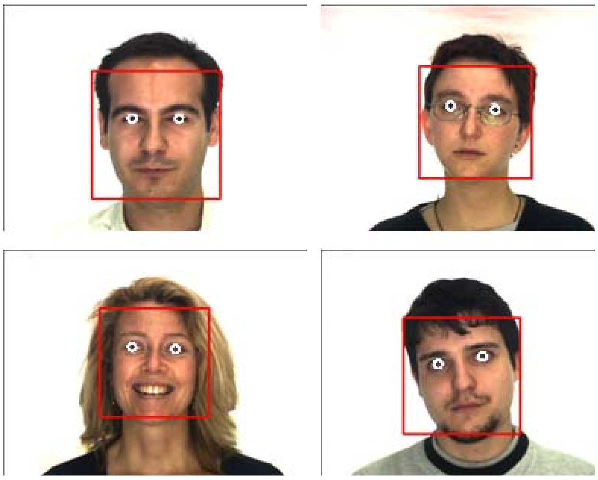

Fig. 4. Results of eye detection on face images from AR database

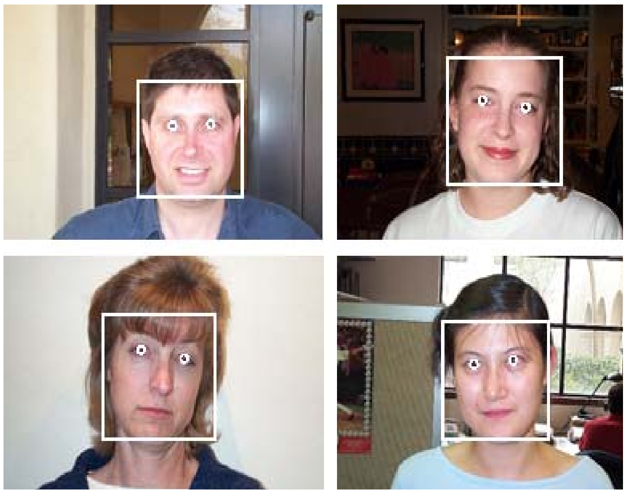

Fig. 5. Results of eye detection on face images from FD database

Dataset	Resolution	# images	Rate	Time
AR	384 × 288	541	91.7%	23ms
FFD	448 × 296	436	92.9%	24ms

Table 1. Eyes detection rate on face area.

2.3. Face registration and photometric normalization

After face and eye had been detected, face and eyes coordinates were obtained. A face can be cropped out of the original image based on the face coordinates and rotated, scaled and shifted based on the eyes coordinates. A point between both eyes was chosen as a reference point. At the very end, an ellipse was pasted over the image to eliminate remaining disturbing factors, such as background or hair (Figure 6). Photometric normalization (Figure 7) was achieved by histogram equalization [10].

2.4. Face recognition and results

System for face recognition has to be trained before it can be used for recognition. A training set with known faces is needed to perform this. Face images from training set have to be registered and photometrically normalized to remove differences among images. For feature selection from face images, an eigenface method [11], [12] was implemented. This method bases on the principal component analysis. When system is trained, it can be used for recognition. Image of unknown face has to be registered, normalized and introduced to the face recognition system. Features were obtained from the image with principal component

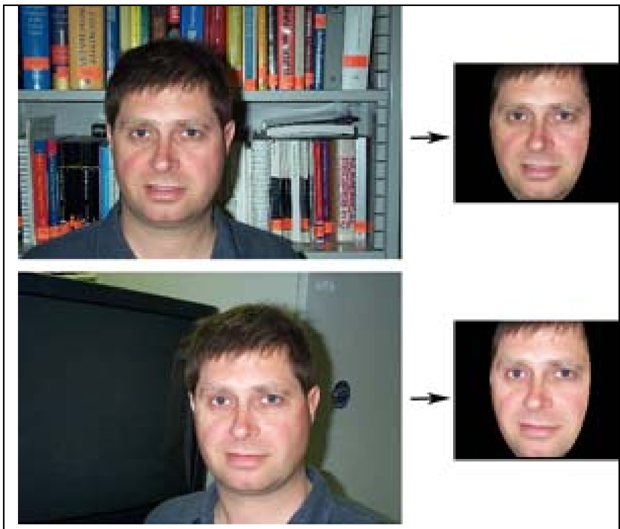

Fig. 6. Results of face image registration.

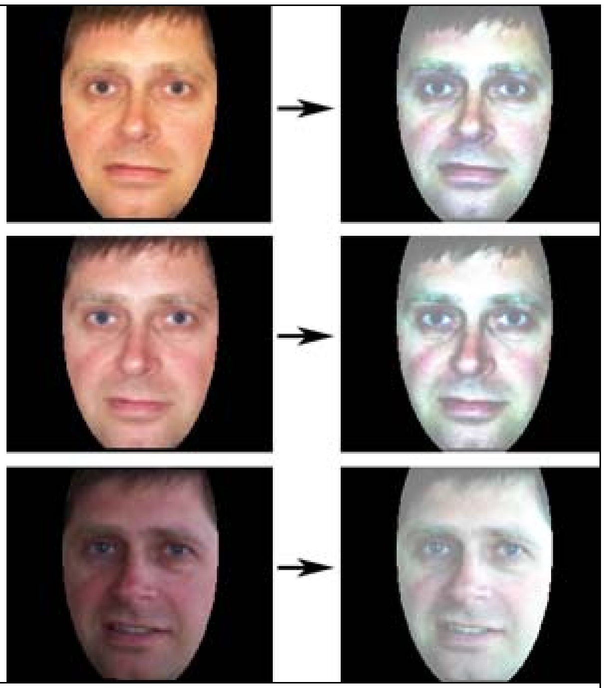

Fig. **7**. Results of face image photometric normalization

analysis and compared to stored features that had been obtained in the training phase. Optimal number of eigenvectors was used in the learning and recognition phase. Results of the described face recognition system are presented bellow. The system was tested on the datasets, which were already mentioned. The results are shown in Table 3. A comparison between registered and non-registered (* in table 3) images was made. This was possible only with the AR dataset, because images from dataset FFD contain reallife backgrounds. A comparison between photometric normalized (EQ in table 3) and non-normalized images is also shown. Face recognition rate was higherwith registered and normalized images as expected and reached 96.2% with FFD dataset. Face recognition system was also tested from the video camera, connected to a personal computer. Around 100 images were recorded for each of the three persons for who the system was trained. The whole procedure including face detection, eye detection, face registration and histogram equalization runs in real-time at 4 images per second.

Dataset	Resolution	# images in training set	# images in testing set	Rate	Rate - EQ
AR*	192 × 244	305	61	68.9%	/
AR	200 × 163	305	61	80.3%	88.5%
FFD	200 × 163	198	157	79.6%	96.2%

Table 2. Face recognition rate.

Face recognition rate was higher with registered and normalized images as expected and reached 96.2% with FFD dataset. Face recognition system was also tested from the video camera, connected to a personal computer. Around 100 images were recorded for each of the three persons for who the system was trained. The whole procedure including face detection, eye detection, face registration and histogramequalization runs in real-time at 4 images per second.

3. BASIC USER MODELING ISSUES

3.1 Personalized content search and retrieval

In previous section we have taken a look at a face detection and recognition techniques. Therefore in second step we have to extend our focus on user profile oriented content search and delivery methods. Today searching methods for particular information (document, image, video,…) usually results in a vast number of hits, with a high amount of irrelevant ones. It is unlikely that some million users are so similar in their interests that one approach to information search fits all needs. Information, content retrieval and delivery can be more effective if individual users' interests and preferences are taken into account. But even the most advanced retrieval techniques cannot prevent avalanches of query results, coming down to the user is important fact. The user would like to have only the most "relevant" query results displayed. To be able to do that, the system must have a mechanism to model users.

3.2 The user interaction mechanisms

The user interaction provide basic observation mechanisms, based on which conclusions about user preferences can be made. Usually these mechanisms are implemented in the framework of the Graphical User Interface (GUI). Related to user interaction mechanisms, users can be either *asked to explicitly* evaluate the suitability of particular content item or the *system has to make implicit conclusions* about the suitability of content based on user reaction (select, ignore, delete, record, user's face mimics, etc.). Combination of both approaches is also used. One would normally expect better results from the explicit feedback approach since implicit feedback systems have to make decisions relying on incomplete and uncertain information. However, some authors report improved results in the domain of television programmes using implicit feedback [13] obtained through the analysis of the PVR usage history.

User models store information about user preferences. They can have a special structure (decision trees, hierarchical structure, keyword vectors etc.) or can contain only lists of content items selected/rejected by the user. The first standardized approach to user modelling in the multimedia (MM) domain is presented in section 4.3.

Content selection and delivery algorithms can be standalone computational procedures (ex. similarity calculations) or can be a part of the user model structure (ex. decision trees).

Generally there are two approaches to content filtering and selection. The first is individual, also known as **content based filtering** (CBF), the other is **collaborative filtering** (CF). The difference between the two is in the process of identification of suitable content for the user. When using the CBF, the suitability of particular content for user is estimated through direct comparison of content description (meta-data) and the model of the active user. When using the CF, the similarity between users is estimated first, and content, liked by users 'similar' to the active user, is recommended. Both approaches have their benefits and drawbacks and can be best used as complementary methods.

3.3 Existing domains of personalized content retrieval

One of the consequences of the Internet "boom" were without a doubt increased research efforts of the adaptive (personalized) hypermedia systems. Apart from personalisation in (hypermedia) documents, personalisation approaches have also been widely used in retrieval and filtering of news (newsgroups and newspapers) and e-mail filtering. While current user modelling techniques are mostly focused on retrieval of textual data items, there are also some examples of personalized retrieval of audio/video and images. Most personalisation efforts have been made, especially in the domain of television broadcasting [17, 18]. They include CB approaches based on textual descriptions; CBF approaches based on comparisons of item ratings between users and hybrid approaches.

3.4. Personalised content selection services within LIVE Production Support System

This chapter gives a quick overview on the two specific content selection services which were identified within the LIVE Production Support System:

1. Content Recommendations (Selection) in the TV Production.
2. Program Recommendations for the individual Consumer.

The LIVE Production Support System (Fig.) consists of several components. The components and user interfaces shown on blue background and outlined with dotted line are associated with user modelling and content selection. The component is named ***LIVE Recommender System***.

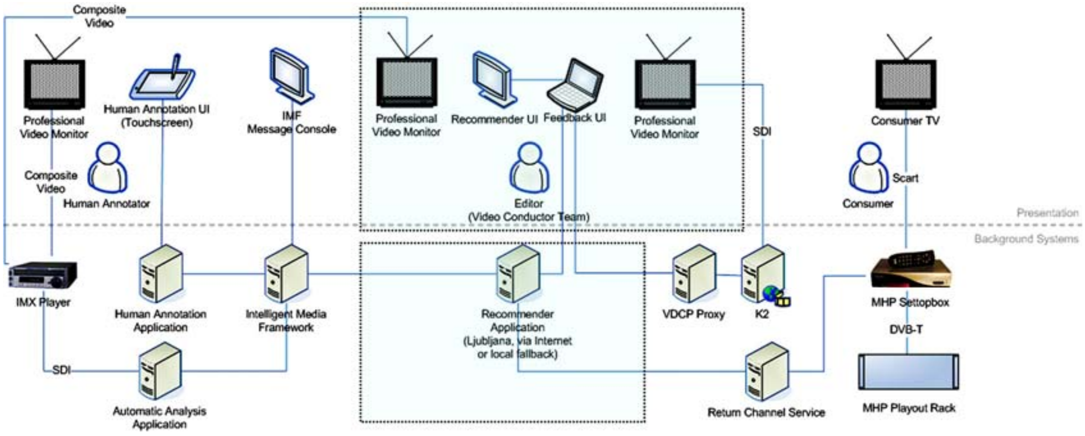

Fig. 8. LIVE Production Support System (First Prototype, copyright IAIS).

The **LIVE Recommender System** component performs tasks related to user modelling, personalisation and content selection services. The tasks which are performed are shown within **Fig.**.

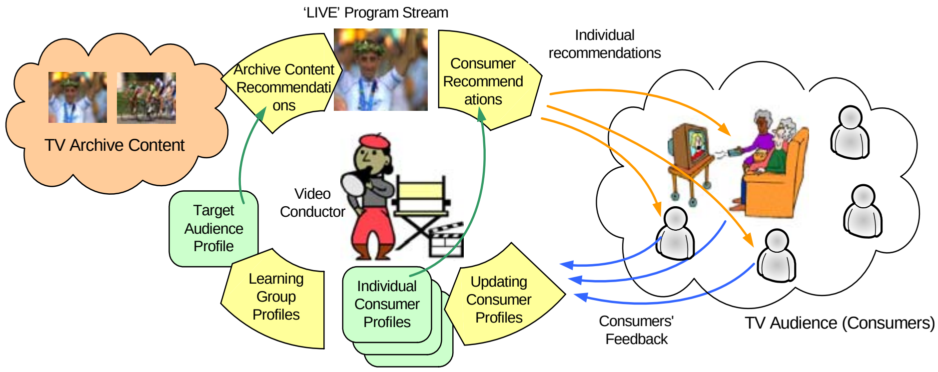

Fig.9: Overview on the tasks of the LIVE Recommender System.

Fig. gives an overview on processes within LIVE Production Support System which are related to content selection and user profiling. The workflow which deals with preparation of content for transmission to the Consumers is as follows.

1. The Video Conductor needs to select appropriate Intelligent Media Assets from the Production Archive to illustrate live sport event or to build parallel stories and channels.
2. The LIVE Recommender System will search the available IMAs and prepare content recommendations based on the selected profile of the target audience. This effectively personalises the content of the TV program according to the target user profile.
3. The "LIVE" TV Program or parts of its content will be recommended to individual Consumers.
4. Consumers will provide feedback on the program they are watching (for example in the form of ratings).
5. Recommender System analyses feedback and computes Consumer and Audience profiles.

From this workflow it is clear that content recommendations (selection of content) can be performed first within the production of the TV program where content from the TV archive is selected, and second it can be used for recommending content to the individual TV Consumers. We can distinguish among the two types of content selection:

- ***Content recommendations for the Video Conductor to support production process***

  A main role of the *Recommender System* is to support new conducting workflows of the LIVE Production Support System by providing automatic recommendations of content suitable to be included in the program, as well as the information on Consumer feedback on the program. The Recommender System implements automatic methods to find and recommend archive content, which will be used to help the Video Conductor in finding the related audio-visual content from the archives. This functionality helps the Video Conductor to react to happenings in live events.

- ***Provision of personalised Consumer content recommendations***

  In the LIVE Production *Support System* support personalised services also for the Consumer. These services will use the Feedback channel, which supports two-way communication, to provide additional personalised information to the single Consumer.
  The Recommender System can provide personalised programme recommendations to the Consumer. According to the Consumer's interests (profile), a personalised list of suitable program items can be compiled and presented to the Consumer. This service is called the **personalised electronic programme guide (pEPG)**. The second variant of the content recommendations for the consumer are the **personalised content alerts**, which are displayed at the Consumer application.

In order to support personalised content selection services, the Production Support System needs a model of the target users. The following chapter describes the role of the user modelling within LIVE Production Support System.

3.5. User modelling

Traditional models of users are based on characteristics of the user which are collectively referred to as cognitive patterns. These patterns include for example: interests, knowledge, preferences, misconceptions, or abilities. Systems incorporating models of user's interest have been widely used to selectively filter information on behalf of users from a large, possibly dynamic information source. A common example of an interest based model is a collaborative filter which infers a user's interest and preferences from the ratings a user applies to an information item and from similarities between user's interests. On the other hand, users might explicitly express their preferences by for example filling a survey.

User model structure. The task of the user modelling will be to define important dimensions of the user model. This structure will be then filled by appropriate values for each of the system users. The structure or dimensions of the model depend on the target application area. The first models within LIVE will be targeted to sport domain, so they will include preferences of the Consumers towards dimensions within sport, such as teams, players, sport disciplines etc. Different user model structure will be needed for professional users.

Context Modelling. In general, the context of something can be thought of as the "extra" (often implicit) information (i.e. associations, facts, assumptions), which makes a full understanding of something and /or an adequate interaction or communication possible. An important issue of context modelling is the decision of which parameters or dimensions are to be taken into account, i.e. modelled. According to the relevant literature, dimensions such as Task, Cognitive Pattern, Relationship, Environment, Granularity, Modality and other can be considered.

Dimensions of the context which will be considered to be included for modelling of LIVE users are the following:

- **Task**: The task of a user is considered important based on the assumption, that the goals of users (who participate in a task) can influence their information needs. When these needs are known in advance, a system can better adapt to its users. Within LIVE, professional users will perform different tasks and thus this information might be considered to influence the response of the system.

- **Environment**: Environmental models influence the interaction between human and computer because they describe the surrounding facts or assumptions which allow a meaningful interpretation of a users' computer usage when their physical environment varies. The Environment dimension might be considered as important within LIVE if the users (the Editor, the Consumer) will require different response based on their environment (for example, the client device or user interface might differ from user to user).

- **Relationship.** This dimension takes into account the interrelationships of individuals. For example, it might be important to know that the group of Consumers is watching a TV program together.

One of the tasks of user modelling within LIVE will be thus to research and define a context information which is needed for providing personalised services.

3.5.1. Role of user modelling research

The basic goal of user modelling research within LIVE project will be to provide models of LIVE users within the system which are needed for the system services and personalisation.
User modelling research will basically consists of offline preparation tasks whose goals will be the following:

- to identify requirements of Production Support System services, what they need to know about users (for example, recommendation service needs to know the preferences of the users),
- to collect and analyse existing information about target users such and store it within Production Support System, such as viewing statistics of the target Consumers,
- to define the structure of the user models, and specify the data structures within the Production Support System to store individual user profiles,
- to analyse user similarities and define user group profiles or user stereotypes,
- to define user modelling and profiling services, which will be performed by the Production Support System. The user profiling services which are needed when the system is online might include: tracking and collecting user's actions (feedback), updating of user profiles through analysis of user feedback, and learning online profiles of the groups of users.

4. FUTURE TRENDS AND EXPECTATIONS

It is clear that user-oriented content delivery systems have a future and that they will appear in most of the content retrieval domains. Regardless of the content types the personalization approaches used today have similar usage scenarios. Normally users browse content listings, implicitly or explicitly rate content items and get content suggestions. The drawbacks of today's systems are context unaware keyword approaches, relatively primitive means of interaction with systems resulting in poor implicit feedback, unawareness of user's mood etc. At this point the question arises: What can we expect in the future from personalized content recommended systems?

4.1.User-system interaction

One of the most important issues, related to the development of personalized systems is the development of digital devices, used for selection and consummation of content. All of these devices from a handheld PDA to the Personal Video Recorder (PVR) have gained in processing power and storage space. They are becoming very similar to computers and can run Java applications, connect to the Internet, download, store and display content items of any type. This means that observation of user preferences and his/her content selections is no longer bound to a single device but is possible in almost any situation at any time: At work, at home on the trip, etc. Consequently, personalized systems will be able to get much more information about the user. This approach will of course require more sophisticated user modeling techniques. Especially difficult issue is acquisition of user preferences from different devices at different times and in different context and their synthesis into a single user model.

Related to this issue is the question of user-system interaction. New types of user interfaces will have to be developed, which will enable seamless communication and interaction. Some authors report that users like to communicate with computers as if they were human; therefore new user interfaces using humanoid avatars and voice processing are being developed. These enhancements may not directly improve the quality of personalized selections, but will nevertheless improve the overall functionality of such systems. Furthermore users prefer communication without too much interaction requesting explicit feedback.

The first step in this direction is the possibility to analyse content usage history in home digital devices, from which some conclusions about user preferences can be made. Another possibility are very simple explicit feedback techniques like "thumbs-up" "thumb-down" used in the TiVo system [20].

Future systems will also be able to track user's focus on the screen of any device and be able to get more reliable information about user's interests. This will be especially useful for analysis of interest in textual documents.

Appropriateness of content selection is very often related to user's current mood. Future systems will be able to identify user's current mood and will recommend him/her suitable music or movie selection for example. Ongoing research projects are investigating the possibilities to identify people's mood based on their biometrical signals (heart rate, temperature, sweating etc.). Complementary initiative comes from the authors of the TV Anytime standard [16], which have in the content description schemes foreseen a field describing the type of mood for which some TV programme is suitable.

4.2. Semantic content search and understanding of content

The problem of contextual understanding of keywords is mostly present in case of textual documents, especially web pages. Current situation in the web is such that data is generally "hard coded" in HTML files. The concept terms used are semantically ambiguous, so there is no way of telling, which of the possible meanings is the right one. For example: a search query "north pole" may provide thousands of result pages with content about the famous geographical location, a company with the same name or even a pub. This ambiguity is transferred into the user profile when information about "preferred" keywords is extracted from web pages and stored. In order to resolve this issue, an initiative called the Semantic web has been started [22]. To put it simply, the idea is to describe specific term meanings, their relationships and context information with the help of schemas and ontologies. The descriptions are made using RDF (Resource Description Format). With the development and use of inference logic this approach may become a very powerful tool for information processing and retrieval.

Similar situation is with MM content types. A solution to this problem is being offered by the MPEG-7 standard. Namely, the MPEG-7 standard aims to describe all types of MM content (audio and speech, moving video, still pictures, graphics and 3D models) including information how objects are combined in scenes, their semantic meanings and relationships (temporal, spatial, etc.). The standard has powerful mechanisms for content description, but is on the other hand very complex, so widespread use is still questionable.

4.3. MPEG-7 user profile

Despite the development in the field of personalized content selection there are surprisingly few standards regarding the descriptions of user preferences. Apart from the widespread use of word vectors there is actually only one standard in this field: the MPEG-7 user profile. The idea was to standardize the description of user preferences and the means of usage of MM content.

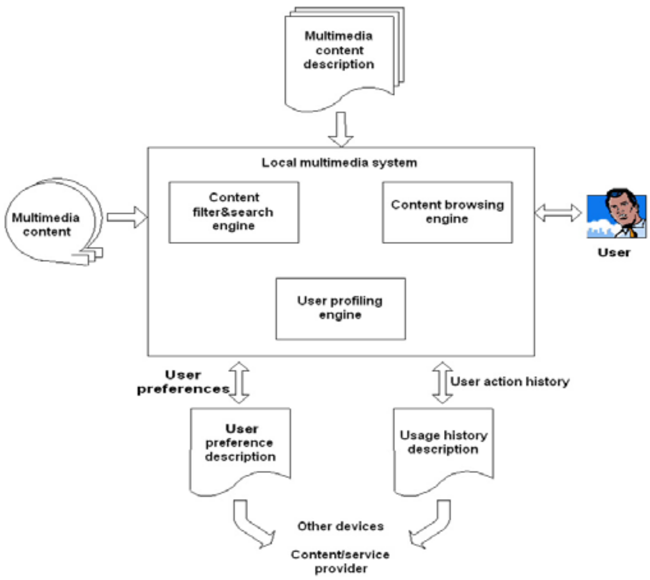

Fig.10. MM system and user interaction

In order to enable the exchange of information about user preferences with 3rd party services they decided to standardize also the information about content usage history. This part of the standard is specifically oriented towards the user interaction with the personal digital videorecorder or similar devices. For this purpose two *description schemes* (DS) were designed: The *UsageHistoryDS* and the *UserPreferencesDS.* The *UsageHistoryDS* enables annotation of user actions (play, record, delete etc.) with respect to particular content item, while the *UserPreferencesDS* enables annotations of user preferences regarding content creation (favourite titles, actors, directors, locations of content creation etc.), content classification (favourite genres, subjects, languages etc.), source preferences (favourite media formats, dissemination mediums etc.) and some others. The standard does not mandate the algorithms used for mapping of usage history to user preferences.

Figure 10 presents typical data and content flow of a personalized retrieval system using user preference description and usage history description.

5. THE ITV CONSUMER AND PERSONALIZED TV APPLICATIONS

The current TV consumer is limited in his interaction with the TV to switching between channels. If he wants to see some specific content, he must put a considerable effort in reading the program guides, remembering the broadcast time and manually switch to target channel. The second option he has is browsing among available channels, which gets inconvenient for large number of channels.

The future iTV consumer will be assisted by **intelligent personalized applications**. These applications will assist him personally to get the right content. They will track the available channels and produce program recommendations according to consumer's profile. In order to provide personalized viewing experience, the additional interactivity levels will be employed to enable:

- switching between subchannels with different viewing angles or different coverage of the same event,
- giving the explicit feedback on the current program by voting or rating,
- setting his personal user profile to enable personalized recommendations such as personalized electronic programme guides
- increasing involvement in the live program by enabling sms or voice comments

An example of the proposed personalized programme guide MHP application is shown in the Figure 11.

Fig.11. The interface of the interactive application showing the personalized electronic program guide.

The consumer will be able to provide the feedback to the production system. The consumer feedback will be an important guide for the live production of the program.

- Explicit user feedback by voting. The user can select if he likes the current program or not. The voting action need to be interpreted in terms of what time section does this voting applies to.
- Implicit user feedback by tracking channels. The watching times of consumers are collected for each subchannel and watching statistics is computed. This is more powerful in combination with consumer profiles.

Fig.12.Voting interface of the consumer application.

The consumer will be able to provide the feedback to the production system. The consumer feedback will be an important guide for the live production of the program.

- Explicit user feedback by voting. The user can select if he likes the current program or not. The voting action need to be interpreted in terms of what time section does this voting applies to.
- Implicit user feedback by tracking channels. The watching times of consumers are collected for each subchannel and watching statistics is computed. This is more powerful in combination with consumer profiles.

5.1. Audience Feedback Interface in LIVE project

With the real-time Audience Feedback tool, the conductor can see different aspects of audience feedback. There can be different feedback statistics for each output stream. The director can see implicit feedback, which can be of demographic nature, for example, nationality, gender or age. There is also an explicit feedback (thumbs up, thumbs down), as a result of an opinion poll, which could be started by the director. The content could also be personalized to the different audience profiles, such as preferences to specific TV program types or different sports. Because of that, the consumer feedback tool should offer also an insight into the number of consumers of different types, which are currently watching the selected content. The consumer feedback tool could also provide overall statistics and trends versus time for each output.

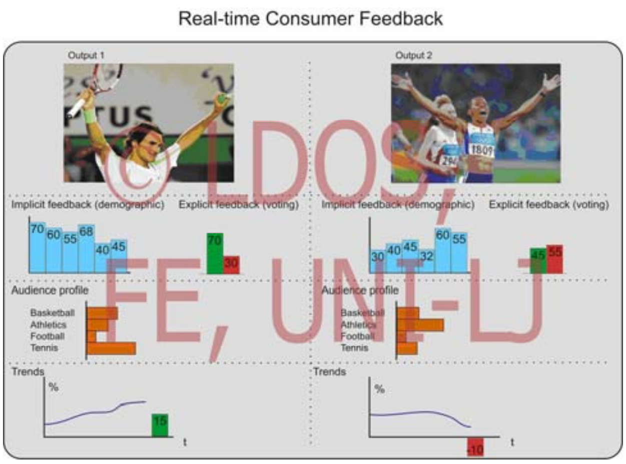

Fig.13. LIVE Audience Feedback interface.

We plan to develop the presented concepts and tools during the LIVE project. This realization will enable a significant step forward in the direction of fully interactive and personalized production of TV channels. Within the project the proposed tools and concepts will be implemented and finally tested during the Olympics 2008 with the real audience. The audience feedback evaluation will give some insights on the suitability of the proposed concepts and the future of the interactive TV production.

6. CONCLUSION

In this paper we presented a possibility of using image processing approach for user authentication and user model based content delivery in TV applications with personalized iTV services. Instead of logging in through the remote control which is not very user friendly, we propose using a set-top box with connected video camera. This system would be possible to process video in real-time and authenticate and track users. As we have seen, face recognition is already successful enough to use it as a way of authentication, especially when using it as an alternative means. We are going to test that kind of user authentication in contrast with classical authentication and feedback collection in the scope of LIVE project. In the future, we are also going to further investigate possibilities of using a video processing system with the personalized iTV services to ease human interaction with devices used in the future interactive services.

The field of personalized content retrieval is also getting a lot of attention. Its results are not only useful but also necessary in order to provide users with efficient tools for content selection. A number of personalization systems exist today, selecting and recommending content to a vast number of users. In the future users can expect improvements in the field of user interfaces, which will enhance the interaction with systems, contextual understanding of terms and topics, enable exchange of user related information and consequently personalized experience on almost any device. This scenario will come true if supporting technologies are provided like unobtrusive biometrical sensors, further advancement of digital devices in terms of processing power and available storage, improved personalization algorithms etc. A very important issue is also usage of standardized content descriptions like MPEG-7 or TV Anytime. Their widespread use is related to development of automatic content indexing methods that include object recognition and consequently semantic understanding of MM content. We should also bear in mind that the data, gathered by personalized systems are very interesting to many commercial companies as well as to the individuals. Therefore we should make sure that these data are not misused and that the user privacy is not compromised.

7. REFERENCES

1. *Hartmann, Producing Interactive Television, Charles River Media, Inc., Hingham, MA, 2002.*
2. *ITV Production Standards Committee, iTV Standards Links,* *http://itvstandards.org/iTVPublic/standards.aspx*
3. *LIVE – Live staging of media events; project web-site: http://www.ist-live.org – Last visited: 10.11.2007*
4. *P. Viola; M. Jones: Rapid object detection using a boosted cascade of simple features. IEEE Conference on Computer Vision and Pattern Recognition, Proceedings,pages 905-910, Hawaii, 2001.*

5. *Tilen Mlakar, Janez Zaletelj, Jurij F. Tasič: Viewer authentication for personalized iTV services, Eighth International Workshop On Image Analysis For Multimedia Interactive Services WIAMIS 2007, Santorini, Greece, 2007*
6. *R. Lienhart; J. P. Maydt: An extended set of haar-likefeatures for rapid object detection. IEEE Conference onImage Processing (ICIP'02), Proceedings, New York,September 2002.*
7. *R-L. Hsu; M. Abdel-Mottaleb; A. K. Jain: Face detection in color images. International Conference onImage Processing 2001, Proceedings, Volume 1, pages 1046-1049, October 2001.*
8. A.M. Martinez; R. Benavente: The AR face database, CVC Tech. Report #24, 1998, http://rvl.www.ecn.purdue.edu/RVL/, accessed: 18.1.2007
9. *M. Weber: Frontal face dataset, California Institute of Technology, http://www.vision.caltech.edu/htmlfiles/ archive.html, accessed: 18.1.2007*
10. *T. Acharya,; A.K. Ray: Image Processing: Principles and Applications, Wiley Interscience, pages 110-115, 2005*
11. *M. A. Turk; A. P. Pentland: Eigenfaces for Recognition. Journal of Cognitive Neuroscience, Vol. 3, No. 1, pages71-86, 1991*
12. *M. A. Turk; A. P. Pentland: Face Recognition Using Eigenfaces. Proceedings of Computer Vision and Pattern Recognition, IEEE, pages 586-591, June 1991*
13. *Gudivada N. et al. Information Retrieval on the World wide web, IEEE Internet Computing, Vol. 1, No. 5, 1997*
14. *Hand D.J. et al. Principles of data mining, The MIT Press, Cambridge, Massachusetts, USA, 2001*
15. *Manjunath B.S. et al. Introduction to MPEG-7, Multimedia description interface, John Wiley & sons Ltd., England, 2002*
16. *TVAnytime: Specification series: S-3 metadata, SP0003 v1.3, part A, 2002 fttp://tva:tva@ftp.bbc.co.uk/pub/Specifications/SP003v13.zip*
17. *O'Sullivan D. et al.: Improving the quality of the personalized electronic program guide, User Modeling and user adapted interaction, Vol.14, 2004*
18. *Cotter P., Smyth B. PTV: Intelligent personalised TV guides, Proceedings of the 12th Innovative Applications of Artificial Intelligence (IAAI-2000) Conference. AAAI Press, 2000*
19. *M. Witkowski et al. Evaluating User Reaction to Character Agent Mediated Displays using Eye-tracking Technology. Workshop on Information Agents in E-commerce; Agents and Cognition, AISB Convention,York, UK, 2001*
20. *TiVo home page, http://www.tivo.com, 2003*
21. *Goulev P., Mamdani E. Investigation of human affect during electronic commerce activities, IEE Eurowearable'03, Birmingham, UK, 2003*
22. *Palmer S., The semantic web: An Introduction, 2001 http://infomesh.net/2001/swintro*
23. *Janez Zaletelj, Tomaž Požrl, Matevž Kunaver, Tilen Mlakar, Marko Meža, Matevž Pogačnik, Andreas Pütz ; Integrated Project FP6 - IST 27312- D6.3: Report on user models and content selection methods;*