Expert Systems with Applications 31 (2006) 164–173
Expert Systems with Applications
www.elsevier.com/locate/eswa

Intelligent broadcasting system and services for personalized semantic contents consumption

Sung Ho Jin, Tae Meon Bae, Yong Man Ro *, Hoi-Rin Kim, Munchurl Kim

Image and Video System Laboratory, Information and Communications University, 119, Munji-ro, Yuseong-gu, Daejeon 305732, South Korea

Abstract

Currently, digital broadcasting tends to personalize the TV watching environment by offering services that can adapt to viewer preferences. In this paper, an intelligent broadcasting system for enhanced personalized-services, based on the semantics of broadcasting contents, is proposed. To implement the system, the MPEG-7 and TV-Anytime Forum (TVAF), as well as agent technology, are employed. For content-level services, real-time content filtering, personalized video skimming, and content-based retrieval using audio characteristic are implemented. To verify the usefulness of the proposed system, we demonstrate it with a test-bed on which content-level personalized services are implemented[1].

Keywords: Personalized broadcasting; Content semantics service; Intelligent system

1. Introduction

Since digital broadcasting was introduced, user-centric broadcasting services have emerged and are increasing. However, current broadcasting technologies support limited services such as reservation recording, simple program guiding with an electronic program guide (EPG) on a personal video recorder system (Tivo), and primitive data broadcasting by broadcasters (Whitaker, 2001).

In order to meet the needs of personalized services with a large amount of broadcasting contents, it is inevitable to understand the semantics of the contents and apply them to broadcasting services. In other words, the new personalized TV watching environment requires the understanding of both the semantics of broadcasting contents and viewer tastes or preferences.

To support user-customized broadcasting services, intelligent broadcasting systems are needed. In order to realize these systems, it is suitable to apply agents to broadcasting systems, which can help interactions between system terminals and give them intellectual functions (Shoham, 1999).

There are currently two main approaches used to develop the user-centric broadcasting environment: one is the system point-of-view and the other is the service point-of-view. The first approach is to employ agent technology to broadcasting environments (FACTS, 1999; FIPA, 2001a; Niiranen, Lugmayr, & Kalli, 2002). In this approach, most works apply the agent technology to broadcasting systems conceptually; to date, there has been little research on practical broadcasting services using the agent.

The other is user-customized broadcasting services (Bais et al., 2002; Ardissono et al., 2003; Rovira et al., 2004). However, for the most part, researchers have tended to center around user preference-based approaches without content-level understanding for recommending contents. This is the limitation of the personalized broadcasting services.

In this paper, a new intelligent broadcasting system and services using content semantics are designed and implemented. To achieve the proposed system, we employed standard technologies such as intelligent agent technology, TVAF, and MPEG-7 audio and visual features, which are used by broadcasting terminals for self-controlling functions and intelligence, for the schema of metadata and user preference, and for analyzing and representing the semantics of the broadcasting contents, respectively. With the proposed system, advanced personalized broadcasting services including real-time content filtering, content-based audio retrieval of broadcasting content and personalized video skimming are demonstrated.

2. Intelligent broadcasting system

Agent platforms in the broadcasting system could provide a user-friendly watching environment by reducing user

* Corresponding author. Tel.: +82 428666129; fax: +82 428666245.
E-mail address: yro@icu.ac.kr (Y.M. Ro).

[1] This work was supported in part by the SmarTV project of the Ministry of Information and Communication, Republic of Korea.

0957-4174/$ - see front matter
doi:10.1016/j.eswa.2005.09.021

operations and adding intellectual functions. We design and implement the proposed intelligent broadcasting system for personalized content consumption. The system overview and functions of system modules are described below.

2.1. System architecture

In a generic broadcasting system with a return channel, a TV viewer needs interactive operations such as the ability to deliver his/her request and find the location of the selected content among various content service providers. In order to minimize viewer efforts in this regard and replace their actions, it is inevitable that the system should have intelligence. The agent which has characteristics of autonomy, intelligence, and co-operability is, therefore, well suited to achieving this objective.

We developed the proposed broadcasting system on top of the FIPA-OS, one of the most popular agent platforms. The proposed system is composed of three independent agent platforms: the User Terminal (UT), which directly interacts with TV viewers; the Service Provider (SP), which responds to a viewer's query and provides metadata or contents to User Terminals; and the Service Information Provider (SIP) which is an intermediary system between User Terminals and Service Providers for exchanging data or necessary information. The architecture of the proposed system is shown in Fig. 1.

This system provides efficient and effective platform interactions through automatic agent network configuration. Active negotiation between the User Terminal and the Service Provider platforms could lead to efficient exchange of content or metadata. Because the agent's message can contain a user's intention or a metadata file, it is unnecessary to assign an additional return channel for interactivity between system terminals. In the system, therefore, it is possible to perform bidirectional broadcasting services using interactivity: for example, the content-based retrieval service between User Terminals and Service Providers or the movement service of user information between User Terminals.

Furthermore, automatic and intelligent management functions (e.g. generation and update) of user preferences are provided by our previous inference method (Kang, Lim, & Kim, 2004), which can support an agent's content recommendations. The following paragraphs provide detailed explanations of the roles of the proposed agent platforms and their agents.

2.1.1. User terminal

The User Terminal is an agent platform embedded in a TV terminal which deals with a viewer's input or output. A viewer can input his/her information or intention and consume desired contents and metadata in this platform. The design goal of the platform is to ensure the viewer-adaptive TV watching environment with broadcasting contents received from a TV receiver.

2.1.2. Service information provider

As previously mentioned, the Service Information Provider interconnects between the User Terminal and the Service Provider. It deals with two databases: one is for the User Terminal's profile, and the other is for the location information of the registered Service Providers. The terminal profile consists of the User Terminal's name and location. It is required when the User Terminal is authenticated for TA

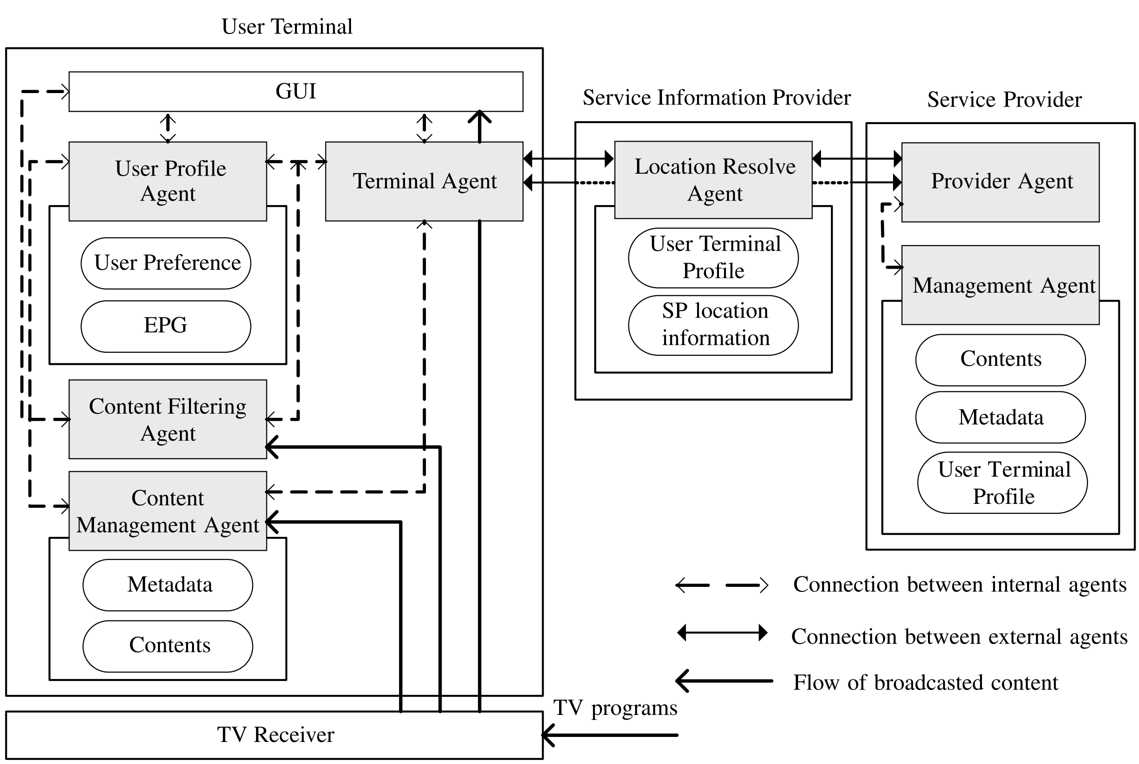

Fig. 1. Architecture of the proposed broadcasting system including three independent agent platforms for personalized broadcasting services.

Table 1
Agent functions in the proposed system

Agent	Function	Platform
UPA	Generating or updating a user's information Offering the user preference and EPG to CMA and CFA, respectively, for personalized content recommendation and content filtering	UT
CMA	Selecting the broadcasting contents coming from a TV receiver depending on the user's tastes, and storing the selected ones Supplying contents or metadata from storage when the UPA demands them for content recommendation	UT
CFA	Filtering live broadcasting contents coming from the receiver in real-time and conveying the results to viewers	UT
TA	Playing the role of a broker between the agents of the UT and the other platforms, e.g., when a user requests a desired content from SPs, TA contacts the SIP to resolve the location information of the providers, and then negotiates with them about the user's request	UT
LRA	Helping the agents of the User Terminals to connect with those of the Service Providers	SIP
PA	Performing similar functions as the TA or the LRA, that is, communicate with the agents of external platforms	SP
MA	Controlling several functions in the Service Provider: storing or updating the database, searching the database for a user's request, and generating video skims	SP

registration. When a user sends a query to the Service Provider, the User Terminal needs the location information to resolve the Service Provider's location.

2.1.3. Service provider

In the proposed broadcasting system, the role of the Service Provider is to respond to user requests with contents or metadata. This platform is designed with two functions: one is to analyze a user's query and offer results matched with the query. The other is to recommend contents or metadata autonomously using a statistical analysis of the user's behaviors.

2.1.4. Agent roles in the proposed platforms

Table 1 shows the role of the defined agents playing inside each platform. Since the agent platform supports protocols for communication between heterogeneous agents (FIPA, 2001b), we develop 'location resolve agent' based on the 'directory facilitator' to provide a yellow pages directory service that is the same as FIPA. Via this agent, an agent can communicate with others registered in the external platform to exchange metadata as well as the user's intention for user-customized services.

Here, the agents are defined as follows: a 'user profile agent' (UPA), a 'terminal agent' (TA), a 'content management agent' (CMA), a 'content filtering agent' (CFA), a 'location resolve agent' (LRA), a 'provider agent' (PA), and a 'management agent' (MA). In addition, UT, SIP, and SP indicate User Terminal, Service Information Provider, and Service Provider, respectively.

2.2. Workflow of system functions

Fig. 2 explains the workflow of agents regarding the system functions in the proposed platforms. Each agent performs their own task depending on the sequences of the workflow. The TA, the LRA, and the PA play central roles in system interaction.

The workflow shown in Fig. 2 depicts content-based broadcasting services using semantics of content which cannot be supported by generic metadata (like EPG). After the Service Provider receives a user's query, the MA performs database searching, finding the original soundtracks for music retrieval or video skimming services. The CFA works for real-time content filtering in which user tastes are analyzed and content semantic information based on user preferences is generated.

The detailed algorithms and procedures for the proposed personalized broadcasting services using content semantics are described in Chapter 3.

2.3. System implementation

We implemented a java-based system using the FIPA-OS. The representation of content metadata follows the schemas of the TVAF metadata, e.g. 'Content Description Metadata', 'Instance Description Metadata', 'Consumer Metadata', and 'Segmentation Metadata' (TVAF, 2001; TVAF, 2002a). And the schema of the user preference used in the system also follows the TVAF standard (TVAF, 2002b; TVAF, 2002c). MPEG-7 audio and visual descriptors are utilized to represent the features of broadcasting contents (Bae et al., 2004; Jin et al., 2004a; Jin, Cho, Bae, & Ro, 2005; Kim, Lee, Nam, Kang, & Ro, 2002). Fig. 3 shows the User Terminal, Service Information Provider, and Service Provider in the proposed broadcasting system. Fig. 3(b) shows the control panel providing personalized services. Its main options are presented in Table 2.

The connection of agents in system terminals is shown in Fig. 3(c). The LRA of the Service Information Provider interlinks two Service Providers and three User Terminals. Here, ① represents the LRA and ② are the PAs of different Service Providers; ③ means the TAs of the User Terminals.

3. Broadcasting services using content semantics

In this section, enhanced personalized-broadcasting services using content semantics are explained in detail.

3.1. Real-time broadcasting content filtering

Let us assume a TV viewer is watching his/her favorite program on a main channel while other channels are broadcasting other interesting programs (e.g. the final round of a soccer game). The viewer would prefer to immediately watch scenes of interest (e.g. shootings and goals) on the other

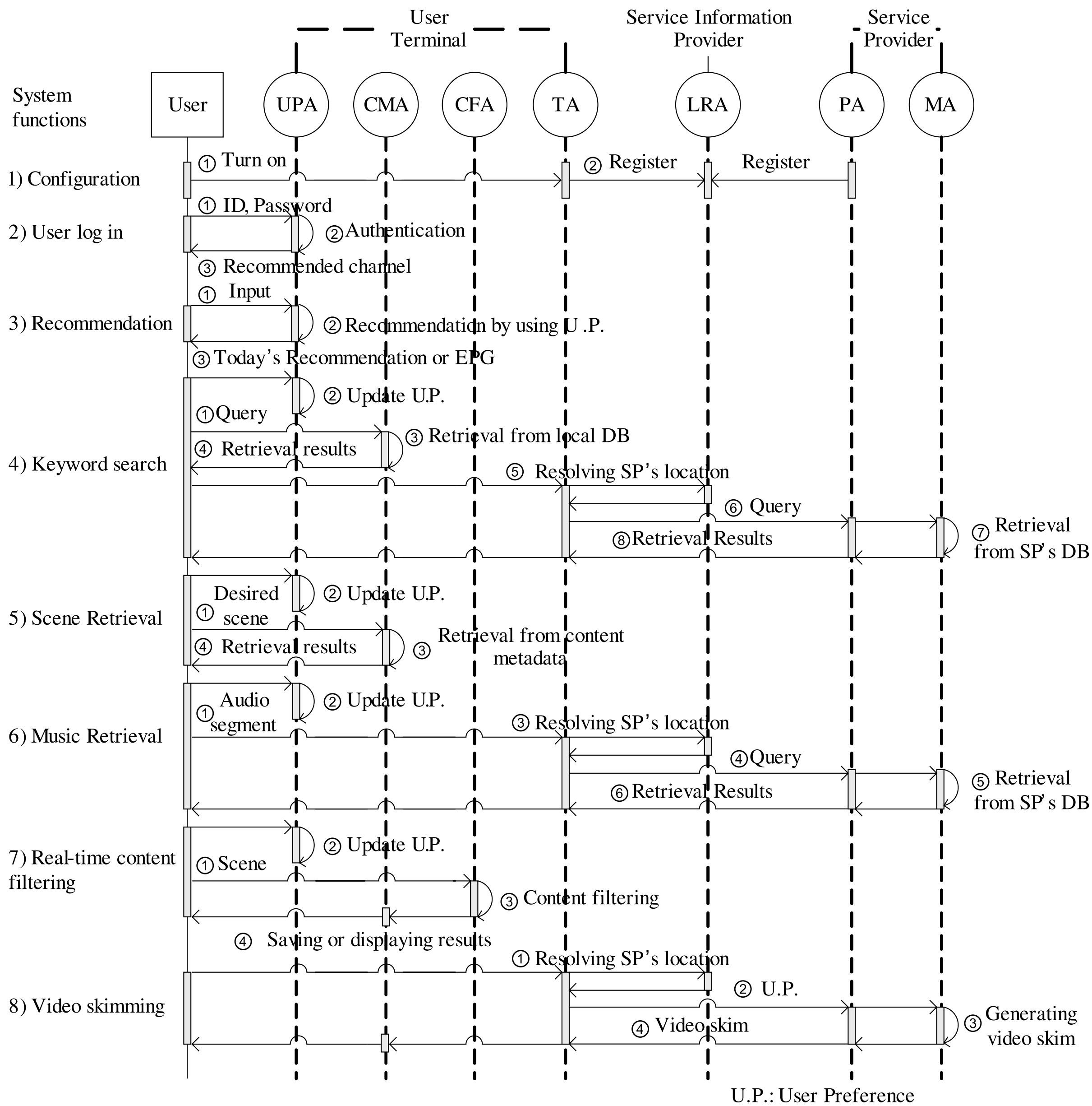

Fig. 2. Sequence diagram for the workflow of representative system functions within the proposed system terminals.

channels. Note that the viewer wants to watch in real-time instead of watching it later after recording the entire soccer game.

If metadata describing program contents does not support content-level indexing or retrieval, the above service scenario is not possible. What is more, a live broadcasting program cannot afford to provide the metadata because it must be made before the time of broadcasting. To achieve the above service scenario, contents-level real-time filtering of live programs should be performed in the TV terminal.

Fig. 4 shows a flowchart of the filtering procedure in which viewers cannot only watch the program of a main channel, but also get semantic information of scenes from other channels of interest.

The procedure of the filtering process can be described in the following steps: (1) A viewer logs-in to his/her terminal. (2) While watching TV, the viewer turns on the real-time filtering function. Then, the viewer can choose filtering options such as choosing a channel of interest, actions after filtering, and desired events. At the same time, the UPA updates one's preference about the preferred events. (3) The CFA controls the filtering process according to the selected options. (4) From the TV receiver, the agent receives video streams of broadcasting contents and samples them with a regular interval. (5) Color features of the sampled frame are extracted in HSI color domain and edge features are extracted to represent spatial transition in the frame. (6) Those features decide the view type of the input frame. (7) Considering the characteristics of temporal frame sequences such as view type transition or continuity, the process detects the preferred event. And (8) finally, the agent gives the matched filtering result to the viewer.

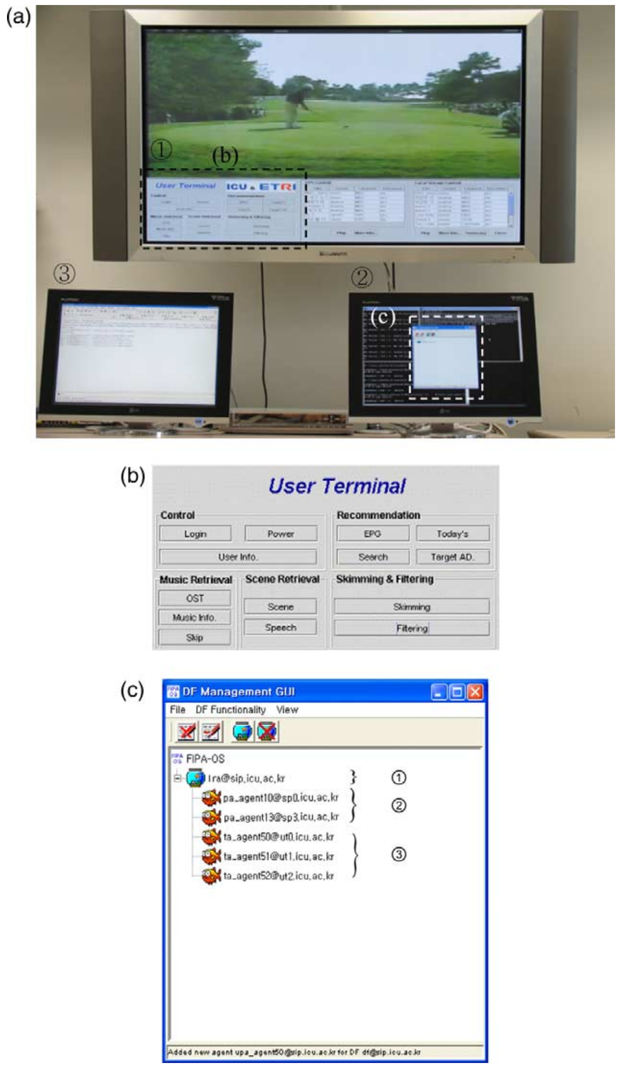

Fig. 3. The proposed broadcasting system applied with agent platforms. (a) Three implemented platforms consisting of ① User Terminal, ② Service Information Provider, and ③ Service Provider; (b) The control panel of the User Terminal which is overlaid on the TV screen; (c) Interface of a 'location resolve agent' in the Service Information Provider connecting with two 'provider agents' and three 'terminal agents'.

In this paper, we implement the filtering algorithm in the sports genre as a prior work. The current filtering algorithm extracts Hue values after transforming the color domain of the sampled video frame (Jin et al., 2005). And, the replay scene can be detected by using the caption in which a scoreboard is displayed. For edge information, Sobel masking, thinning, and the Hough transform are employed.

As shown in Fig. 5, a viewer can choose the filtering options for real-time content filtering, where the channel of interest is displayed in the box on the bottom right-hand side of the figure. As soon as the content filtering agent captures a viewer's desired event, the terminal notifies the viewer of the fact by means of a channel change or channel interrupt functions.

Currently, our filtering algorithm shows close to 80% filtering performance in regard to 'Shoot' and 'Goal' scenes in 6 h of soccer game videos in real-time. The scene decision and content filtering are made with 3 frames/s of the frame sampling rate at a TV terminal with 650 MHz CPU, which has similar capability with the set-top box currently being used (Jin et al., 2005).

3.2. Personalized video skimming

TV viewers have individual tastes about semantic scenes or events of broadcasting contents. The goal of this service is to increase viewer satisfaction with broadcasting content. In other words, the viewer can be served with different video trailers of the same content depending on their preferences. Service providers or content makers can promote the consumption of their contents by the personalized video skimming service. Additionally, it provides facilities for TV viewers to select their favorite channels from multiple channels.

Fig. 6 depicts the procedure for the proposed user-customized video skimming service. We assume that this service is a subscriber-based one because of the security of user preferences. Under this assumption, the procedure of the service can be described as follows: (1) While a viewer watches TV, one sends his/her user preference to a Service Provider for an appropriate video skim. At this time, both a TA and a PA work for exchanging data such as the preference and the video skim. (2) The MA of the provider parses the received preference to know what kind of events the user desires. (3) If the pre-segmented video clips about the corresponding events exist within a database, the agent makes a video skim by using them. Otherwise, the agent generates a video skim with the video clips defined by the provider. (4) Then, the provider delivers the generated skim to the corresponding user terminal and arouses the user's interest about the content. And (5) the UPA of the User Terminal recommends the received video skim.

Fig. 7 shows the implementation of the skimming service in the User Terminal. In this experiment, a woman logs-in to this terminal and sends her preference to one service provider in which an EPG shows that a movie is scheduled for

Table 2
Main options in control panel

Option	Function
Control	A user can register his/her information and log-in or –out
Recommendation	It offers user preference-based functions in which a user can see the content lists of EPG or local storage filtered by agents, search contents by text-keywords, and etc
Music retrieval	Its function is to search original soundtracks or similar music based on contents
Scene retrieval	Using metadata from Service Providers, it allows users to skip and see meaningful scenes in broadcasting contents
Skimming and filtering	It runs personalized video skimming and real-time content filtering

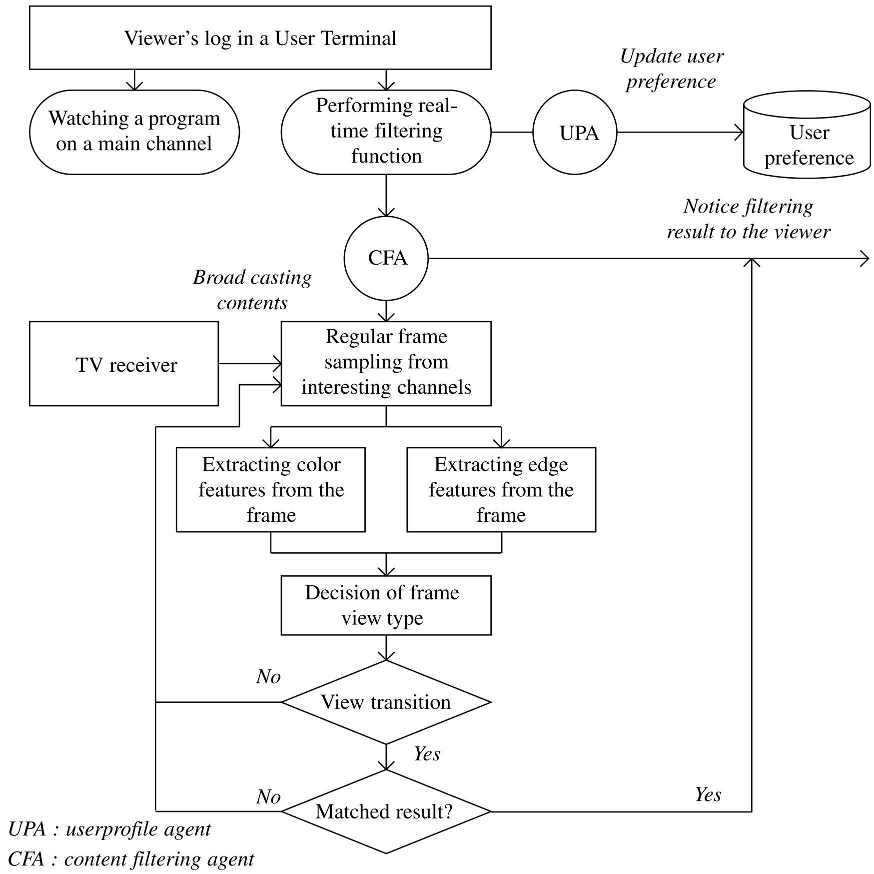

Fig. 4. Flowchart for the real-time broadcasting content filtering service in the proposed broadcasting system. It can have multi-channels as channels of interest according to system requirements.

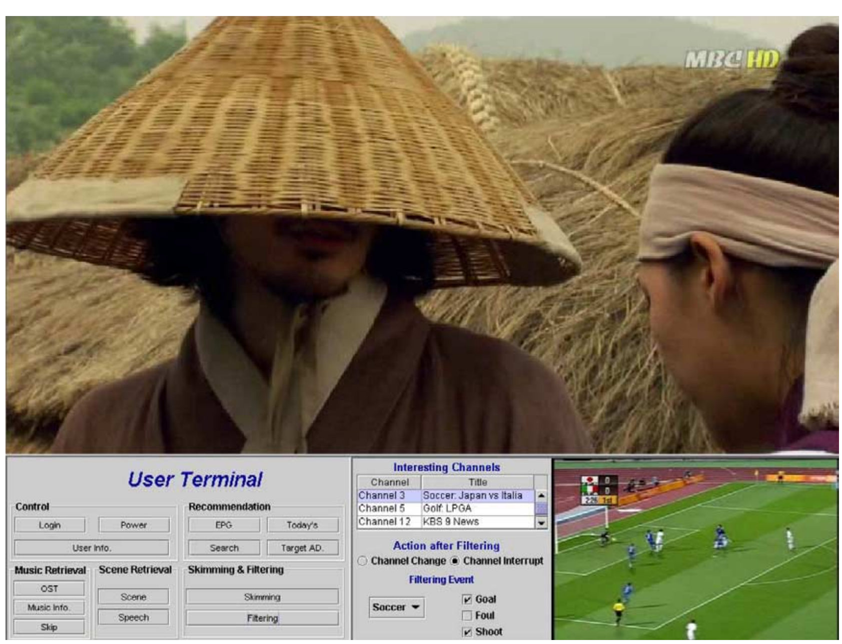

Fig. 5. The screen shot to run the real-time content filtering service with a single channel of interest (in this case, a soccer game). The box on the bottom right-hand side of the screen in the above figure displays the desired broadcasting content from the channel of interest. This log-in user selects a soccer game on Channel 3 as his/her channel of interest, and chooses 'Shoot' and 'Goal' scenes as the desired semantic information. On filtering them, the terminal interrupts the main screen by displaying the semantic video segment of the selected channel of interest.

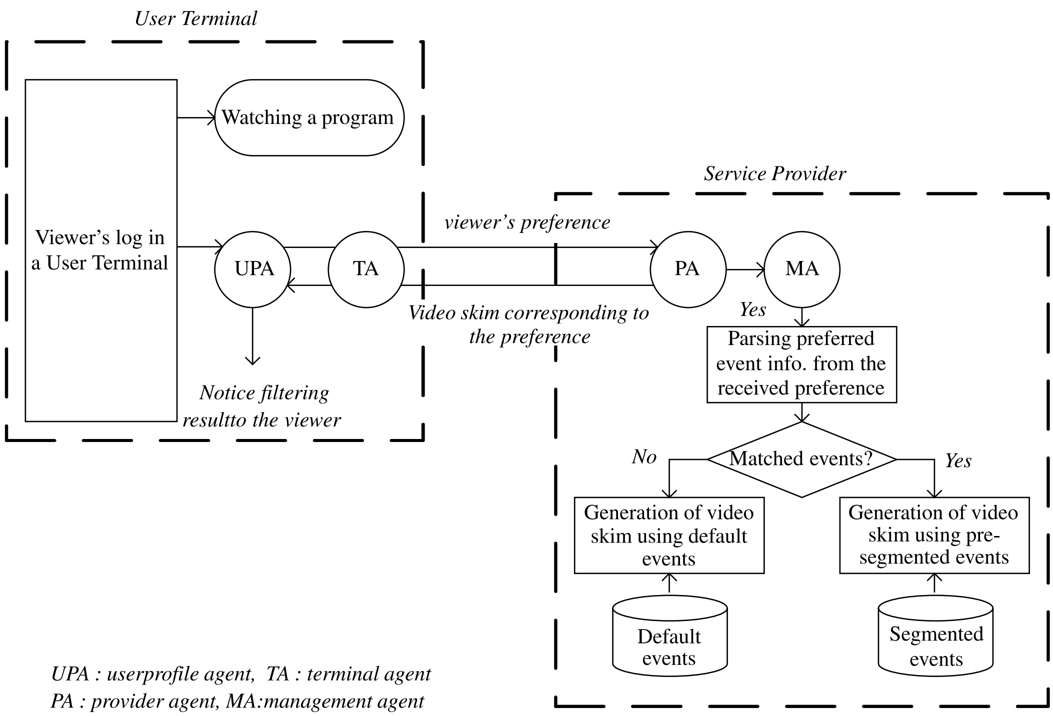

Fig. 6. Flowchart of the personalized video skim service in the proposed broadcasting system.

broadcasting. Because she enjoys 'Romantic' scenes, the small display shows the corresponding skim as a movie preview that is mainly composed of the 'Romanic' scenes.

To establish this service, we need to extend the schema of the user preference. As previously mentioned, the main schema is from the TVAF standard. But because there is no element to express a user's preferred events, we insert an 'Events' element in the "ClassificationPreferences" elements as show in Fig. 8. The Service Provider parses this element to generate a video skim.

3.3. Content-based retrieval using audio characteristics

Without pre-made metadata, the proposed system serves TV-viewers with content-based retrieval using audio characteristics of a broadcasted content. The goal of this service is to search similar music or original soundtracks (OST) to meet the viewer's request. The function for similar music not only provides the matched music, but also retrieves similar music in three categories such as timbre, rhythm, and mood. The OST retrieval provides the music chosen from the broadcasting content in high quality, and also offers auxiliary information such as title, singer, and album name. The significant disparity between similar music and OST retrievals is in the size of the database. OST retrieval can be performed in a music database consisting of about 20 songs. In contrast to the former, similar music retrieval should have a large amount of music database.

In the implemented Service Provider, the retrieval process uses a hybrid method using both pitch histogram and MFCC-VQ dynamic patterns which contain both static patterns and temporal patterns of melody.

Fig. 9 represents a flowchart of the music retrieval service based on audio characteristics which are shown in the following procedure: (1) A viewer watches a broadcasting content containing OST or music. (2) According to the viewer's intention, a TV terminal sends the selected audio segment information of the content to a Service Provider. Here,

Fig. 7. The screen shot to run the personalized video skimming service. The status window shows that this viewer shows interest in the 'Romantic' scenes in the movie genre. Therefore, the box on the bottom right-hand side in the above figure displays a video skim as a movie trailer which is made out of several 'Romantic' scenes of the movie.

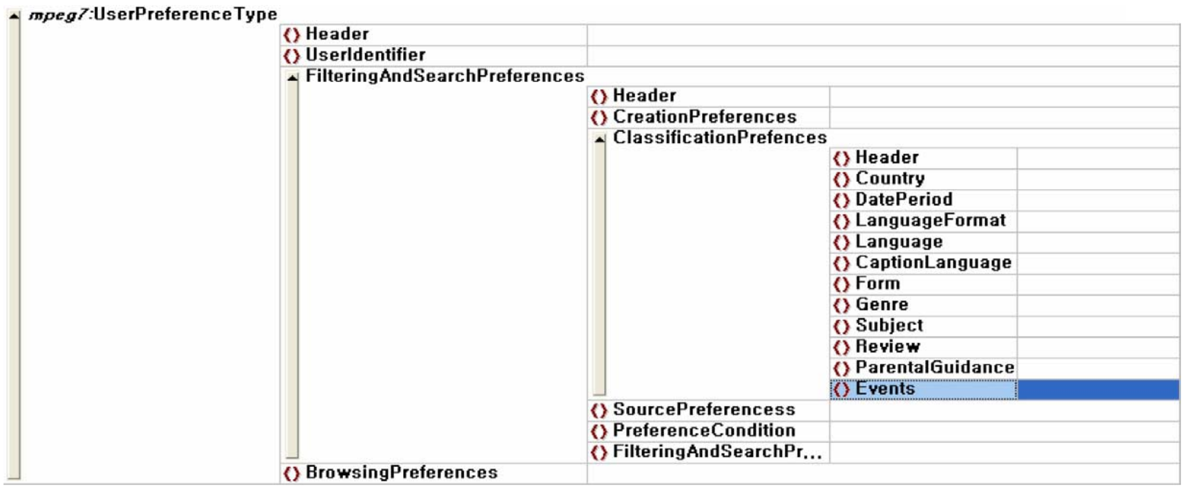

Fig. 8. Enhanced grid view of the XML schema of the extended user preference showing the inserted 'Events' element.

the audio segment information includes the time position in which the user's request meets the content. And (3) using the delivered segment information, the MA extracts audio features of the MPEG-7 Fundamental Frequency descriptor and MFCC from the corresponding content segment, and performs a matching process within its music database. After that, the matched results including auxiliary data are replied to the viewer (Park et al., 2004).

As a result of this audio service, a viewer can listen to clear music without noise such as actors' speeches, background effects sound, and etc. In particular, similar music retrieval acts like an on-line jukebox, and is located on the bottom right-hand side of the screen in Fig. 10(a). For similar music retrieval, the music database is composed of more than one thousand songs which are extracted from compact discs. By using the hybrid method and two-step ordering, the performance shows about 90% retrieval accuracy for OST, and about 76.3% retrieval accuracy for similar music.

As mentioned at the beginning of this section, we also implement generic metadata services including text keyword-based retrieval, semantic video scene retrieval using metadata, user information mobility, content recommendation based on

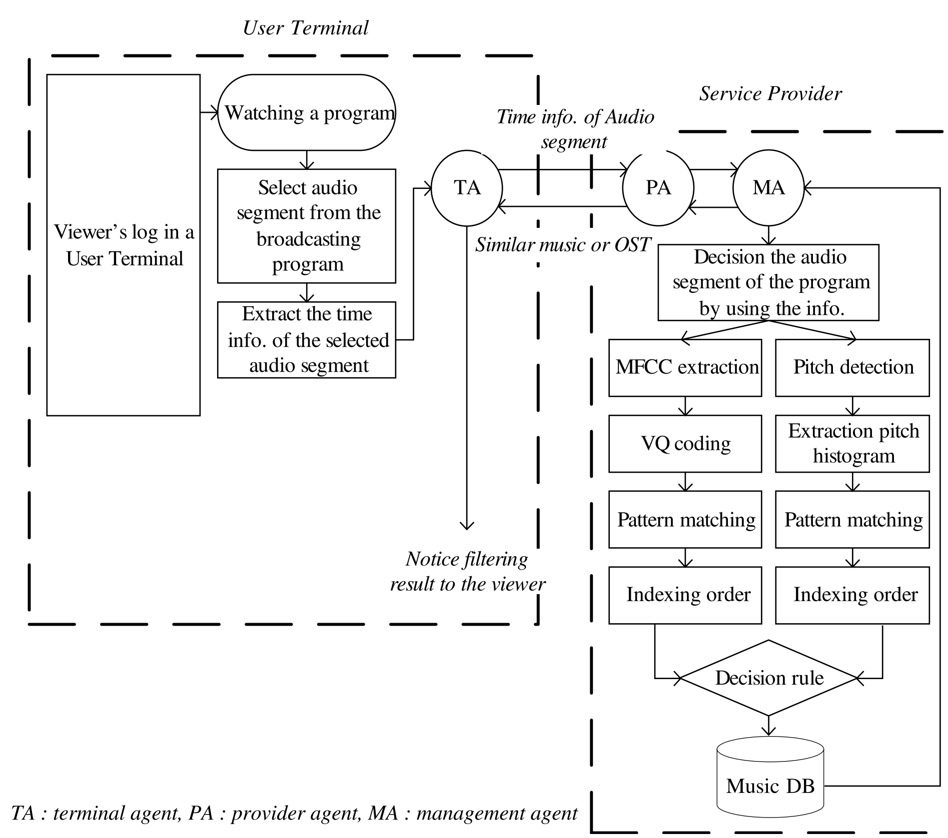

Fig. 9. Flowchart for the content-based retrieval service using audio characteristics of a broadcasting content in the proposed system.

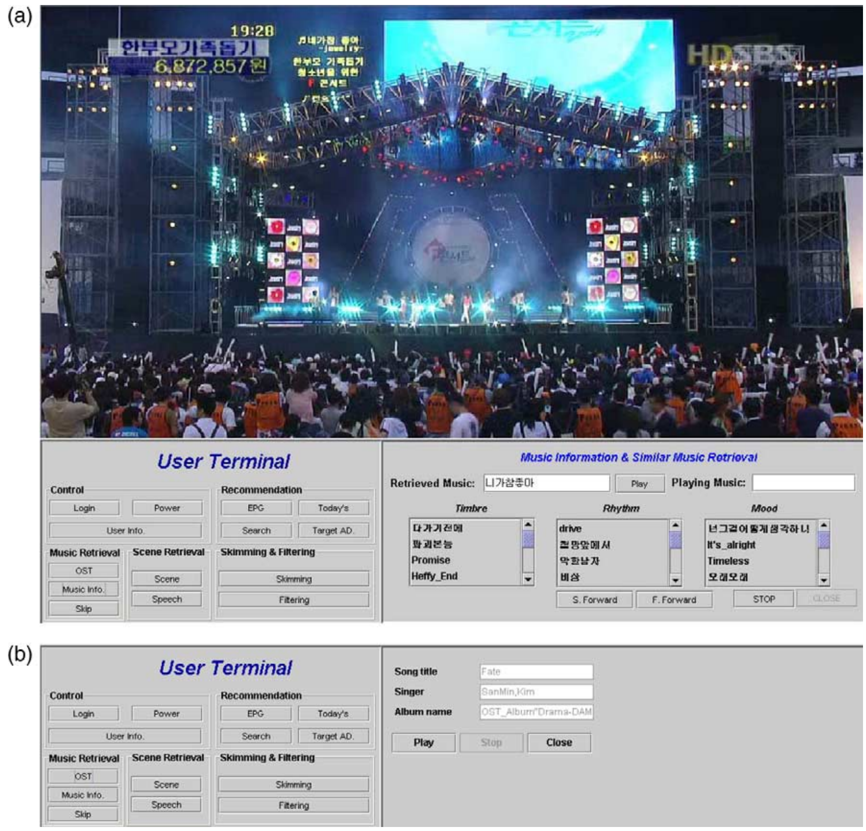

Fig. 10. The screen shots to run the content-based retrieval service using audio characteristics of a broadcasted content. (a) By the audio characteristics of pop music obtained from a music program, similar music in three categories—timbre, rhythm, and mood—are retrieved. Each category shows similar songs with 10 ranks; (b) As the execution result of the OST segment chosen from the 'DAMO' drama, the OST retrieval not only plays corresponding music, 'Fate', in high quality, but also gives auxiliary information such as song title, singer name, and album name.

user preference, target advertisement by inference of gender and age, and etc (Jin et al., 2004b).

4. Discussion

In this paper, we have attempted to form new TV terminals and services on an intelligent broadcasting system. By applying intelligent agent technology, the interactivity between system platforms was achieved in the proposed broadcasting system. This allows for the efficient and effective exchange of user requests and provider responses. Another important point of our work is that it practically employs user tastes to content semantics, not a content list. The services using the semantics provide a more user-friendly TV watching environment.

When, for example, an interesting sports game is being broadcasted, because of its characteristics, it is highly probable that it will be applied to the filtering service whose aim is to filter broadcasting programs in real-time. Although the current filtering algorithm is structured around a few sports genres, it shows very high performance in those genres and establishes a representative service model using content semantics. The OST or similar music retrieval on the broadcasting content shows the possibility of extension to the domain of T-commerce in which TV viewers can purchase desired music from service providers. Personalized video skimming is expected to support benefits for both users and service providers and is also very practical.

Despite these results and accomplishments, there is still much to be studied about the performance enhancement of the used algorithms, particularly in terms of information security and the integration on a set-top box. We expect that these problems could be solved through future works.

5. Conclusion

In this paper, we proposed a new broadcasting system and user-customized services based on intelligent agent technology. Moreover, for applying content semantics to the services, MPEG-7 and TVAF standards are used. In order to demonstrate the usefulness of the new broadcasting environment, we implemented a test-bed which had several agent platforms over the IP network, and experimented with three services including real-time content filtering, personalized video skimming, and content-based retrieval using audio characteristics. Proceeding from what has been said above, it should be concluded that the proposed agent system and services offer a new broadcasting environment which is more user-friendly and provides greater facilities for TV viewers.

Acknowledgements

We would like to thank Mr Sungtak Kim, Mr Jun Ho Cho, and Mr Mun Jo Kim who participated in this project.

References

Ardissono, L., Gena, C., Torasso, P., Bellifemine, F., Chiarotto, A., Difino, A., et al. (2003). *Personalized recommendation of TV programs Lecture notes in artificial intelligence (Vol. 2829)* (pp. 474–486).

Bae, T. M., Jin, S. H., Choo, J. H., Park, M., Ro, Y. M., & Kim, H., et al. (2004). Video Segmentation by Hidden Markov Model Using Multimodal MPEG-7 descriptors. *Proceedings of SPIE2004* (*Vol. 5304*) (pp. 214–223). San Jose.

Bais, M., Cosmas, J., Dosch, C., Engelsberg, A., Erk, A., Hansen, P. S., et al. (2002). Customized television: Standards compliant advanced digital television. *IEEE Transactions on Broadcasting*, *48*(2), 151–158.

FACTS (1999). *AC317 FACTS work package a2 public deliverable phase 1 deliverable part 1-project overview*. AC317/NN/WP2/DS/P/011/b1.

FIPA (2001a). *FIPA audio/visual entertainment and broadcasting specification*. XC00081B.

FIPA (2001b). *FIPA abstract architecture specification*. XC00001J.

Jin, S. H., Bae, T. M., Choo, J. H., & Ro, Y.M. (2004). Video genre classification using multimodal features. *Proceedings of SPIE2004*, San Jose (*Vol. 5307*) (pp. 307–318).

Jin, S. H., Bae, T. M., & Ro, Y. M. (2004). Intelligent agent-based system for personalized broadcasting services. *Proceedings of CISST2004*, Las Vegas (pp. 613–619).

Jin, S. H., Cho, J. H., Bae, T. M., & Ro, Y. M. (2005). Content filtering for broadcasting media. *Proceedings of SPIE2005*, San Jose (*Vol. 5670*) (pp. 260–271).

Kang, S., Lim, J., & Kim, M. (2004). Modeling the user preference on broadcasting contents using Bayesian networks. *Proceeding of SPIE 2004*, San Jose (*Vol. 5308*) (pp. 958–967).

Kim, C. S., Lee, H. K., Nam, J., Kang, K., & Ro, Y. M. (2002). Semantic event detection using MPEG-7 descriptors for interactive broadcasting. *Proceedings of CISST2002*, Las Vegas (*Vol. 1*) (pp. 391–395).

Niiranen, S., Lugmayr, A., & Kalli, S. (2002). Agent-based personalization in digital television. *Proceedings of 5th Nordic Signal Processing Symposium, Trondheim.*

Park, M., Park, C., Kim, H., Im, S., Shin, D., & Yang, S. (2004). Content-based music information retrieval using pitch histogram of band pass filter signal. *Proceedings of AIRS2004*, Beijing (pp. 245–248).

Rovira M., Gonzdlez J., Lopez A., Mas J., Puig A., Fabregat J., et al., (2004) IndexTV: A MPEG-7 based personalized recommendation system for digital TV. *IEEE international conference on multimedia and expo, Kobe* (pp. 823–826).

Shoham, Y. (1999). What we talk about when we talk about software agents. *IEEE Intelligent Systems*, *14*(2), 28–31.

Tivo, http://www.tivo.com.

TVAF (2001). *Specification series: S-2 on system description, SP002v1.3*.

TVAF (2002). *Specification series: S-3 on metadata part A: Metadata schemes, SP003v1.2*.

TVAF (2002). *Exchanging TV-anytime metadata over IP networks, AN462*.

TVAF (2002). *Specification series: S-6 on delivery of metadata over a bi-directional network, SP006v1.0*.

Whitaker, J. (2001). *Interactive television demystified*. New York, NY: McGraw-Hill.