PERSONALIZATION OF HYPERLINKED VIDEO IN INTERACTIVE TELEVISION

Benoit Huet, Joakim Jiten, Bernard Merialdo

Institut EURECOM, Sophia Antipolis, France
{huet, jiten ,merialdo}@eurecom.fr

ABSTRACT

The GMF4iTV project is building an end-to-end broadcast system for providing interactivity to TV programs through active video objects. This paper describes the personalization scheme that has been defined and implemented in this project. On the Authoring side, an interface tool allows the video producers to annotate both objects and associated additional content, with attributes from different ontologies. On the receiver side, a personalization engine compares these attributes with the user profile to personalize the interaction with the user. We describe the various components of the system and provide some application examples.

1. INTRODUCTION

Although the interest for hypervideo has already been demonstrated in several projects, its usage remains limited, in great part because of the lack of appropriate platforms for creation, distribution and presentation. For example, the Hypercafé project [1] is a nice illustration and application of a hypervideo engine. The Viper system [7] allows creating personalized programs through the selection of clips at the user side. Other systems like Hyper-Hitchcock [11] link video clips through various level of details. The GMF4iTV system [10] proposes a complete solution for hypervideo in the context of interactive TV. In this paper, we present the personalization components of the GMF4iTV system, based on ontologies and first order logic rules. There is growing interest in the use of ontologies and inference engines [8] [9] as a mean to provide more flexible decision making mechanism. Our experiments demonstrate the use of such powerful tools for the purpose of hyperlinked video personalization.

2. THE GMF4ITV SYSTEM

The objective of the GMF4iTV project is to develop a complete end-to-end platform to produce and broadcast Interactive TV Programs where interaction is performed at the video object level. GMF4iTV relies on an upgraded version of the MHP (Multimedia Home Platform) software at the receiver end (set-top box or PC). In the Authoring phase, a regular TV program is edited to define video objects: objects are selected on one video frame and tracked throughout the corresponding video shot (see fig.1). The additional content (which may be text, graphics, HTML page, MPEG-4 clip or MHP application) is associated to each object.

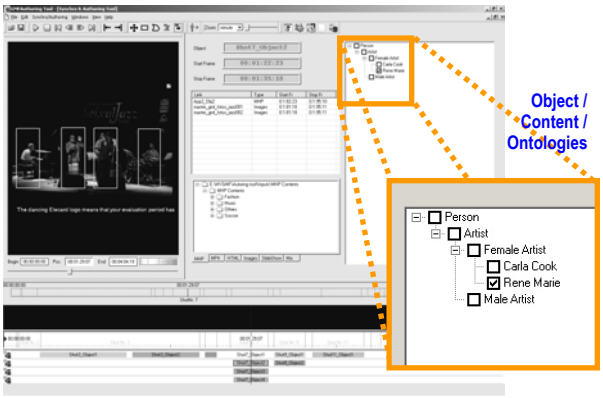

Fig 1: Example of the Authoring Tool

Then, the additional information (objects and their location, additional content) is encoded in MPEG-7, and multiplexed with the original MPEG-2 Video stream into a single transport stream. The combined signal is sent to a receiver running the MHP Reference Implementation software. The receiver de-multiplexes the transport stream to retrieve the original information, which is stored in local storage until required for user interaction. The information about object location is used by the video player to superpose a visual indicator (for example a colored bounding box) over the video object. The user may select visible objects using a regular remote control, or a PDA. When an object is selected, the associated additional content is displayed to the user.
When interaction is performed with the remote control, the bounding boxes and the additional content appear on the regular TV screen. This might be disturbing if several people are watching at the same time. Therefore, the system allows the possibility of interaction through a PDA. A MPEG-4 version of the MPEG-2 program is created at production time, and sent within the program

0-7803-9332-5/05/$20.00 ©2005 IEEE

stream. The PDA interacts with the Set-Top box through a regular WiFi connexion. In this case, the bounding boxes and the additional content appear on the PDA only, so that the user may conduct his interaction without disturbing other viewers.

On the production side, the steps involved in the preparation and broadcast of an interactive program are the following:

- Authoring,
- Video object definition (manual selection, followed by automatic tracking),
- Association of additional content to objects,
- Metadata encoding in MPEG-7,
- Synchronization (to insure that additional content will be entirely transmitted when the object appears),
- Multiplexing with the original MPEG-2 video,
- Receiver side,
- De-multiplexing of metadata, local storage of additional content,
- Display of video object bounding boxes,
- When selected by the user, presentation of additional content associated to the object.

Personalization in GMF4iTV

Personalization is the process of adapting the behavior of the system to the user. In GMF4iTV, this happens at two different steps:

- Object selection: the authoring phase defines a number of objects in the video as active objects. Based on a user profile, personalization will activate or de-activate these objects, so that only objects of potential interest are highlighted.
- Additional content selection: an object might have several pieces of additional content associated (called versions). Personalization will select which version should be displayed to this user.

Authoring an interactive program is a time-consuming process, even when semi-automatic tools are used to speed the operation. Furthermore, the total bandwidth of the broadcast channel is limited, therefore the amount of additional material that can be transmitted with the program is constrained. For those reasons, we envision that personalization in our context will be used to define broad categories of users, or adapt the interaction to the current mind of users, and not for a fine profiling of the viewers.

We now detail how the two Personalization functions are implemented in the complete system.

3. AUTHORING AND PERSONALIZATION

Ontologies

In order to personalize the behavior of objects and additional content, it is necessary to identify some semantic characteristics of these items. This is done by annotating these items with attributes during the authoring phase. Rather than considering a raw list of possible attributes, we have decided to organize the possible attributes in an ontology. An ontology is a tree structure where more specific concepts are children of more general concepts. The annotation is performed by selecting the item (object or additional content) and then choosing the adequate node in the tree in the annotation interface as depicted in Fig. 1.

The main advantage of using ontology is that the selection rules for the objects can use general concepts which cover several more specific object types. Therefore, when a program has been annotated with the greater level of detail, changing the personalization behavior will only require changing the decision rules without re-annotating the data. This gives a maximum flexibility to the definition of personalization.

Another advantage is that, in the case where a limited amount of time is available, the annotation can be done at a coarser level, without losing the consistency with other programs.

For each program, two ontologies are defined: one for objects, one for additional content. The ontology for additional content has a flat structure (a root with a single level of children), as there was never a need for more complexity in our scenarios. The object ontology is more complex, depending on the scenario. Right now, each program has its own object ontology, but in the future, we envision that there should be one ontology per program genre, so that the ontology creation and maintenance is not part of the annotating process.

Static User profile

The static user profile resides at the receiver side. It contains general user information (not related to any specific program), and its content might be diverse: age, sex, geographical location, preferences… This profile is considered as an external component, which is updated by the end-user.

For example, the user profile might contain the following information:

```
(user sex male)
(user age 40)
```

Rules

The personalization process takes decisions by comparing the object attributes with some of the user characteristics. These decisions can be of very complex type, and should be easily constructed and modified when personalization is defined for a program. Therefore, we have decided to use an inference engine to implement these decisions. The decision rules are just data for this engine. For a given program (or eventually program genre), it is needed to express the personalization decisions in this rule language. This allows to define rules which are not simple

comparisons between attributes, but which can perform complex deductions (such as generalization through the ontology for example).
Below are two examples of rules appearing in a Music Show program, one to activate information about musicians, one to display concert information in the adequate geographical region, based on user location:

If (user selected artist_info) and (isa obj Artist) then activate obj
If (isa user_department Region1) and (isa concert Region1) then activate concert

Transient Profile
While the static user profile contains persistent information about the user, there are two cases where this information may not be sufficient:
• When a program requires a user information which is not available in the profile, for example whether the user is interested in whales or not.
• When some preferences are time dependent, and they should be adapted to the user mood at the time where the program will be viewed.
To handle these cases, we have added the possibility of a transient user profile, which is a set of questions which are asked to the user at the beginning of the program. This allows the user to enter the relevant information at the time when the program is viewed.
For example, the interaction may generate the following transient profile:

(user select lyrics)

MPEG-7 Encoding
All the personalization information: ontologies, rules, object and additional content attributes, is encoded in an MPEG-7 structure and embedded in the GMF4iTV Metadata [3] [4] . This allows coherence with the other components of the GMF4iTV system.

4. METADATA TRANSMISSION

After the authoring phase, the following metadata is available besides the original MPEG-2 video:
• A MPEG-4 translation of the MPEG-2 video for display on the PDA,
• The list of video objects and their positions (frame and location) in the video program,
• The additional content associated to the objects (this might be pictures, text, html, MPEG-4, or even MHP application),
• The personalization information (ontologies, rules, and attribute annotations).

This data is multiplexed in a single MPEG-2 transport stream. As the total bandwidth is limited, the metadata cannot be transmitted instantaneously. Therefore, synchronization algorithm has been developed to schedule the transmission of the metadata and to insure that the metadata associated to an object has been entirely transmitted before the object appears in the video [2] .

5. PERSONALIZATION ENGINE

On the receiver side, the set-top box de-multiplexes the MPEG-2 stream to recover the different components, and activates the TV display and the Object Interaction manager. This OI manager activates the personalization engine at the beginning of each program. The initialization procedure first locates all facts and rules from the de-multiplexed MPEG-7 document thanks to GMF4iTV specific tags and asserts them together with the first-order statements contained in the locally stored user profile. Subsequently, the inference engine applies the rules in order to assert all possible new facts.

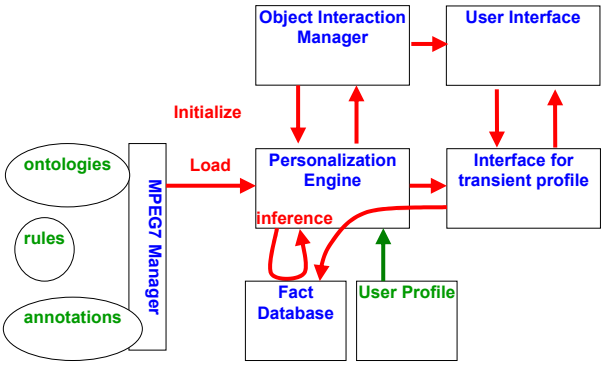

Fig 2: Role of the Personalization Engine

Once the engine initialization is completed, the Object Interaction manager is ready to make decisions about whether or not objects should be highlighted and which extra content should be displayed if selected by the user using its remote control.

6. APPLICATION EXAMPLES

Several scenarios have been developed, in cooperation with video producers, to show the potential applications of hyperlinked video in interactive TV programs [10] . Some of these scenarios include the use of personalization, which we utilize to illustrate the previously discussed ideas.

Personalization of Object Selection
The purpose of object personalization is to highlight video objects that are interesting to the viewer. As mentioned earlier, this is done by an inference engine that matches facts about objects with the viewer's preferences. The fashion show scenario illustrates this idea. A fashion show

is a program where models walk along a scene and present different types of clothes to the public. We define the moving objects as the models, which have a personalization attribute of either male or female. The user profile indicates whether the user is male or female, and the rules of the personalization engine indicate that male (resp. female) models should be activated if the user is male (resp. female).

Personalization of additional content

The extra content associated with an object is annotated with a version label, describing its nature. When the user selects an highlighted object, the personalization engine decides which version of the additional content is displayed. In the Documentary scenario, the moving objects are animals, and the associated content consists in either detailed explanations on their living habits for adults or a quiz game (MHP application) for children. Therefore, through their PDAs, several viewers may have different interactions with the same TV program.

Fig 3: Illustration of different personalized user interactions from the same program

Advanced Personalization Scenarios

While the previous examples are simplistic in their decision scheme, we also have developed more advanced scenarios to illustrate the flexibility and power of the decision rules:

The music program shows musicians playing a sequence of songs. The additional content consists of the lyrics of the songs, information about the artists' discography and several lists of their concerts to come, organized by geographical region. The additional content is activated by user selection. In the case of the lists of concerts, the user profile contains geographical information (the city name or zip code where the user lives) and the program ontology contains the necessary knowledge for the inference engine to match the user zip code to the correct geographic regions. Then, the personalization engine is able to select the appropriate regional list to display to the user.

7. CONCLUSION

In this paper, we have described the personalization capabilities for interactive TV programs of the GMF4iTV system. The system contains a complete production chain, from the identification and annotation of objects, link to additional content, semantic annotation, MPEG-7 encoding, multiplexing with the MPEG-2 video stream, decoding on the set-top box and interaction with the user through a PDA. Personalization is possible both for object activation and additional content selection. The personalization process is based on rules which are processed by an inference engine running on the set-top box. This powerful process allows a very flexible usage of personalization in a variety of situations provided by different application scenarios. A public demonstration of the entire project was given at the IBC 2004 Conference.

9. ACKNOWLEDGEMENTS

The work presented here is a component of the GMF4iTV Project. The complete system is built thanks to the contributions of all project partners. This project is funded by the European Union under contract IST-2001-34861

10. REFERENCES

[1] Sawhney, N.; Balcom, D.; Smith, I.; Authoring and navigating video in space and time, IEEE Multimedia, Vol 4 , Issue 4 , Oct.-Dec. 1997, Pages:30 – 39

[2] López, D. Gonzàlez, J. Fabregat, A. Puig, J. Mas, M. Noé, E Villalón, F. Enrich and G. Fernàndez : Synchronized MPEG-7 metadata broadcasting over DVB networks in an MHP application framework, IBC'2003, Amsterdam, September 2003.

[3] Nack, F., Lindsay A.: Everything You Want to Know About MPEG-7: Part I & Part II, IEEE Multimedia, Vol. 6 No. 3 & 4, (July - September 1999) 65-77, (October - December 1999) 64-73

[4] MPEG-7 library: http://iis.joanneum.at/MPEG-7

[5] Dai, W., Cohen, R. (2003). Dynamic Personalized TV Recommendation System. InProceedings of the 3rd Workshop on Personalization in TV. Pittsburgh, U.S.A.

[6] H. Lee, J. Nam, B. Bae, M. Kim, K. Kang and J. Kim: Personalized Contents Guide and Browsing based on User Preference. Second International Conference on Adaptive Hypermedia and Adaptive Web Based Systems: Workshop on Personalization in Future TV, Malaga, Spain, 2001, pp.131-150.

[7] Stefan Agamanolis and V. Michael Bove, Jr., Viper: A Framework for Responsive Television, IEEE MultiMedia, vol. 10, no. 1, July - September 2003, pp. 88 – 98

[8] I. Horrocks et al.: The Ontology Inference Layer OIL, Technical Report, Vrije Universiteit Amsterdam, June 2000

[9] M. A. Musen, et al.: Component-Based Support for Building Knowledge-Acquisition Systems. Conference on Intelligent Information Processing, August 2000, pp 18-22.

[10] B. Cardoso, et al.: Hyperlinked video with moving objects in Digital Television, ICME 2005, Amsterdam, July 2005.

[11] Andreas Girgensohn, Frank Shipman, and Lynn Wilcox: Hyper-Hitchcock: Authoring Interactive Videos and Generating Interactive Summaries Proc. ACM Multimedia 2003. pp. 92-93, November 2, 2003.