SPATION: Services Platforms and Applications for Transparent Information management in an in-hOme Network

Mauro Barbieri*, Gerhard Mekenkamp*,
Nicola adami**, Riccardo Leonardi**, Pierangelo Migliorati**

*Philips Research, Prof. Holstlaan 4, 5656 AA, Eindhoven, The Netherlands, +31 40 27 44491, {Gerhard.mekenkamp, mauro.barbieri@philips.com}

**University of Brescia, Via Branze 38, 2523, Brescia, Italy, {adami, leon, pier@ing.unibs.it}

Context and objectives

The characteristics of PCs, huge storage capacity, tremendous processing power, and high flexibility are becoming available for consumer devices like set-top boxes, TVs, and VCRs. Interconnection of these devices and wireless communication with various portable devices will create a complex home system with the capacity to store many types of data and offer new ways of interacting with it. To offer the user high flexibility and ease of use, new solutions are required. Advanced retrieval methods are needed to support accessing data stored anywhere in the home system from any device. Meta-data obtained through analysis, services, and logging user behaviour is needed to support these functions. Transfer of data must be easy, and transfer and adaptation of accompanying meta-data must be transparent to the user. The combination of broadcast, storage, and internet will open the way to new types of applications and interactions with the home system.

A large distributed storage space will be available in future home networks consisting of CE equipment, PCs and handheld devices. The objective of the project is to find innovative solutions for the movement, organization and retrieval of information in such a heterogeneous home system.
Three major technical issues are under consideration:
1) New Meta-data computing methods are needed to support advanced retrieval methods. This means ways to solve how to generate meta-data by analysing the content, how to combine meta-data from various sources and how to transform meta-data for use by different devices.
2) New services providing meta-data, applications and UIs to make retrieval of information easier for non-IT-expert users.
3) Standards for inter-storage communication need to be extended in the area of handheld devices, meta-data storage and services.

Metadata Generation, Combination and Transformation

Several methods for automatic **generation** of audiovisual indexes will be presented and demonstrated. With respect the state of the art in the multimedia indexing area the attention has been paid to develop algorithms for the automatic generation of video previews and pictorial video summaries.

The problem of metadata **combination** and **transformation** has been studied in MPEG-7 and MPEG-21 context. Two different aspects have been considered: the integration of MPEG-7 descriptions referring to the same multimedia documents and the transformation of MPEG-7 description by metadata filtering and scaling.
The goal of MPEG-21 is to define a multimedia framework in order to allow the easy and transparent access of any multimedia resource from any terminal in the network. In the use case here described, the attention has been placed on the word "easy". This means that, because of the limited visualization capabilities of the handheld, some adaptation of the content-descriptions has to be performed. For example, when a user wants explore the content of a video using the k-frames associated to the relative shots, at least the k-frame size and the color representation are adapted in order to fit the display characteristics of the handheld.

However, in general, more than one description associated to the same document is available; for example, multiple descriptions of the same content obtained by different extraction methods. In this case, the user can choose a particular description reflecting his/her characteristics or can select the integration of all available descriptions. In the second case, the descriptions must be adapted in order to obtain a unique and, generally, richer description with respect to the initial ones.

Advanced Application for Multimedia Browsing and Retrieval

In Spation demonstrators are being developed with the aim of investigating ways to find and manage content in home networks in as simple a manner as possible.

The SPATION demonstrator consists of several prototypes of stationary CE devices (e.g. TV, PVR, Audio jukebox, photo browser, etc.) and handheld devices interconnected in a home network. Each prototype is implemented with a common PC augmented with extra hardware boards such as tuners and mpeg decoders.

The demo shows a personal video recorder capable of automatically generating audio-visual indexes of MPEG-2 digital video broadcast streams. A video browsing application running on a handheld device allows users to browse and search in a large home video archive. Three different modalities for content-based video navigation are demonstrated: browsing by means of representative k-frames, video previews and video pictorial overviews.
The video browser allows navigating through thumbnails of key-frames representing meaningful scenes of a selected program. Key-frame selection is based on clustering the results of a simple but effective shot-cut detection algorithm that runs during video recording.

When a large amount of video content is available it is difficult for a user to find something interesting. A method that helps the users to decide whether a certain recorded program is interesting enough to watch is provided by automatically generated video previews. A video preview (or *video skim*) is a short video clip, for example 90 seconds, that provides an overview of a recorded program. During recording, content analysis algorithms compute low-level descriptions of the content [1] [2] such as scene boundaries, scene complexity, motion activity, human presence, audio volume, etc. When the recording is complete and all the content descriptions are available, a number of heuristic rules are applied to select the scenes that are most relevant for users that want to preview the content without actually spoiling it. Examples of such rules are non-disclosure of the end of a program, inclusion of action scenes and dialogues, etc. Concatenating the most relevant scenes generates video previews that convey key-aspects of a video program and helps users judge whether or not they would like to see the entire content. The actual video preview is stored as metadata in the form of list of pointers to the parts of the original content that are included in the preview. Users can watch directly video previews on TV sets or stream transcoded versions to handheld devices.

Using a *pictorial overview* of a recorded video program, users can quickly see what a program is about, remember whether he or she has seen it before or decide whether or not to watch it or to download it to a portable device. The pictorial overview is a set of the 15 most relevant pictures extracted from the original content. An automatic summarization algorithm [3] selects a set of images that best represents a particular piece of content.

This demo, not only shows the feasibility of the proposed algorithms will be shown, but also how different types of generated summaries can be used in future CE-devices.

Acknowledgements

References

[1] N. Dimitrova, S. Jeannin, J. Nesvadba, T. McGee, L. Agnihotri, G. Mekenkamp, "Real time commercial detection using MPEG features", 9th Int. Conf. On Information Processing and Management of Uncertainty in Knowledge-Based Systems, July 2002, Annecy, France.
[2] Barbieri M., Mekenkamp G., Huet B., Yahiaoui I., Merialdo B., Leonardi R., Rose M., "Generating TV Summaries for CE-devices", *Proc. 2002 ACM Int. Conf. On Multimedia*, Juan les Pins, France, December 2002.
[3] Huet B., Yahiaoui I., Merialdo B., "Optimal video summaries for simulated evaluation", *Proc. of the European Workshop con Content-Based Multimedia Indexing*, Brescia, Italy, September 2001.