Prototype Demonstration: Video Content Personalization for IPTV Services

David C. Gibbon, Zhu Liu, Harris Drucker,
Bernard Renger, Lee Begeja, and Behzad Shahraray
AT&T Labs – Research
New Jersey, USA
{dcg,zliu,renger,lee,behzad}@research.att.com
drucker@monmouth.edu

1. Introduction

IPTV customers will have access to thousands of video content sources and will require powerful yet intuitive tools to locate desired content. We propose a solution based on stored user interest profiles and multimodal processing for content segmentation to produce manageable content subsets for users. Segmentation involves linguistic processing of extracted closed caption text combined with video shot boundary detection as well as audio processing. Query relevance ranking with temporal and other metadata constraints is used to form timely, focused content sets for users.

We will demonstrate a prototype for browsing automatically personalized video content in a set-top usage context, which meets the challenging requirements imposed by the low resolution TV display and the limited input capabilities of a simple infrared remote control. The client runs on Microsoft Windows Media Center Edition (MCE) and the video content is derived from MPEG-2 DVR content that has been indexed and transcoded to WM9 to facilitate streaming from a Windows Media Server. The MCE client application interfaces to the AT&T Labs – Research MIRACLE platform [1] which creates the personalized content from a 36,000 program archive that is updated on a daily basis from a wide range of content sources.

Content personalization has been an active research area for a number of years. Dimitrova et al, have developed systems for personalizing DVR content with similar design goals in mind [2]. We have reported on content personalization and repurposing for mobile devices [3] and this demonstration extends that work to support a set-top box usage scenario.

2. User interaction

New users register with the prototype service and establish an initial profile of interests using a traditional form-filling web application. Multiple topics of interest can be specified using sets of search terms and content sources for each topic.

Thereafter, the users interact with the service using a set-top box terminal and conventional Infrared (IR) remote control. Users first identify themselves to the system (Figure 1.) It is envisioned that this self identification step can be omitted in certain cases such as for a single user household. Next, a personalized content synopsis for the user's first topic of interest is displayed to the user (Figure 2.) The display includes scrolling processed closed caption text to simulate the roll-up captioning mode and still images represent the clips for rapid visual identification. Using the arrow keys on the IR remote, users can select different topics, and for each topic, different clips to view. Once selected, the video plays in a PiP (picture-in-picture) window which can be toggled to full screen playback. The content synopsis display is updated several times daily as new content is acquired by the service and deemed to be of interest to the user.

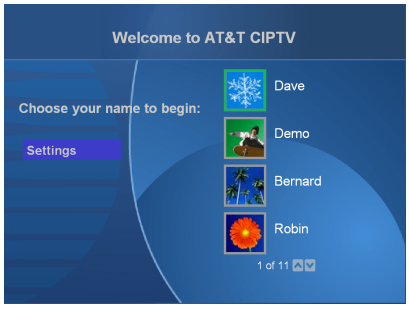

Figure 1**:** User identification on an IPTV set-top box

3. System architecture

Acquisition/processing: An acquisition subsystem consists of a set of DVRs each with 200GB local storage, hardware MPEG-2 encoders and IP networking for the EPG data service. A system status monitoring and content management component controls the flow of media files to a processing subsystem. The processing subsystem extracts global metadata, generates content indices using multimedia processing and transcodes the video into WM9. Finally, the media and index information are published to a large (6TB) content archive and HTTP application server and the personalization subsystem is informed that new content is available.

The processing includes a content segmentation module which logically breaks long-form video (30 to 90 minutes) into topical clips (1 to 10 minutes) based on multimodal story segmentation. Further details of the media processing are described in [1].

Personalization Module: User profiles contain sets of queries represented in XML format. For each user topic, a set of relevant clips is derived from the database based on query term matches, temporal and other metadata constraints. The process follows that described in [3] but here we use pre-segmentation instead of query-based segmentation.

Figure 2: Synopsis display for a particular user's topics of interest

4. Demonstration Components

Figure 3 is a pictorial representation of the demonstration system. An XBOX 360 with a universal media remote control drives an SD or HDTV monitor. This is connected to a PC or laptop running MCE which can also run the demonstration using an LCD monitor at 1024x768 resolution and a USB IR receiver. The user interface is generated by an HTTP application server, and a media server delivers the streaming video to an embedded video player on the MCE. This is the same player that is used for playing back live or recorded television, so consistency of the user experience is maintained.

Figure 4 depicts the software architecture for the prototype which is based on available consumer electronics equipment and leverages standards-compliant components such as XSL, XML, HTTP, etc. where possible.

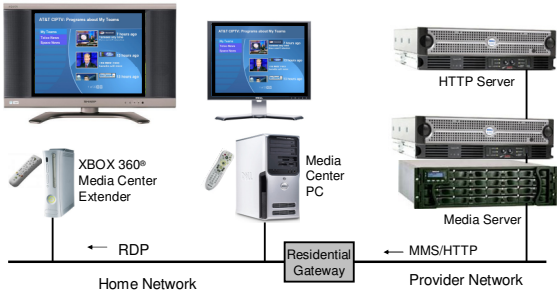

Figure 3: Prototype system components

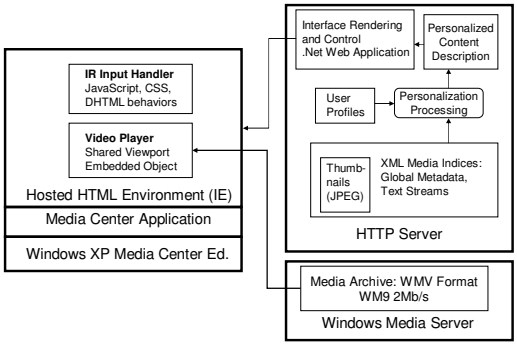

Figure 4: Prototype Software Architecture

5. References

[1] Z. Liu, D. Gibbon, B. Shahraray, "Multimedia Content Acquisition and Processing in the MIRACLE System," in *IEEE CCNC,* January 7, 2006.

[2] The Video Scout System: Content-based analysis and retrieval for Personal Video Recorders, N. Dimitrova, R. S. Jasinschi, L. Agnihotri, J. Zimmerman, T. McGee, D. Li, in Handbook of Video Databases, CRC Press, June 2003, ISBN 0-8493-7006-X.

[3] D. Gibbon, L. Begeja, Z. Liu, B. Renger, and B. Shahraray, "Creating Personalized Video Presentations using Multimodal Processing," Handbook of Multimedia Databases, Edited by Borko Furht, CRC Press, pp. 1107-1131, June 2003, ISBN 0-8493-7006-X.