

Multimodal Interaction with Speech, Gestures and Haptic Feedback in a Media Center Application

Markku Turunen¹, Jaakko Hakulinen¹, Juho Hella¹, Juha-Pekka Rajaniemi¹, Alekski Melto¹, Erno Mäkinen¹, Jussi Rantala¹, Tomi Heimonen¹, Tuuli Laivo¹, Hannu Soronen², Mervi Hansen², Pellervo Valkama¹, Toni Miettinen¹, and Roope Raisamo¹

¹ University of Tampere, Tampere Unit for Computer-Human Interaction, Finland
{firstname.surname@cs.uta.fi}

² Tampere University of Technology, The Unit of Human-Centered Technology, Finland
{firstname.surname@tut.fi}

Abstract. We demonstrate interaction with a multimodal media center application. Mobile phone-based interface includes speech and gesture input and haptic feedback. The setup resembles our long-term public pilot study, where a living room environment containing the application was constructed inside a local media museum allowing visitors to freely test the system.

Keywords: Speech, haptics, gestures, multimodal interaction, media center.

1 Media Center Application and its Multimodal Interface

This demonstration shows how speech, gestures, and haptic feedback can provide an efficient and accessible interface for a media center for different user groups (regular users, blind users, visually impaired users, and physically impaired users). The media center application [1] is based on a low-cost PC (Athlon X2 3800) running the media center server software, a mobile device (Nokia N95) for interacting with the user and running the client software, a wireless access point to connect these together, and a high-definition 40" digital television displaying the user interface. The server software implements the media center application functionality and includes speech recognition and speech synthesis. The mobile phone has embedded gesture recognizer, speech recognizer, haptic feedback controller, speech synthesizer, and GUI. This setup is illustrated in Figure 1.

The media center application offers the same functionality as the other applications and set-top boxes in this context, e.g., users are able to watch and record television broadcasts, listen music, and view photographs. Currently, it provides full control over digital television content, including an advanced electronic program guide (EPG). As illustrated in Figure 1, the system has two graphical user interfaces. A large television screen acts as the main GUI, while the mobile phone GUI provides additional information.

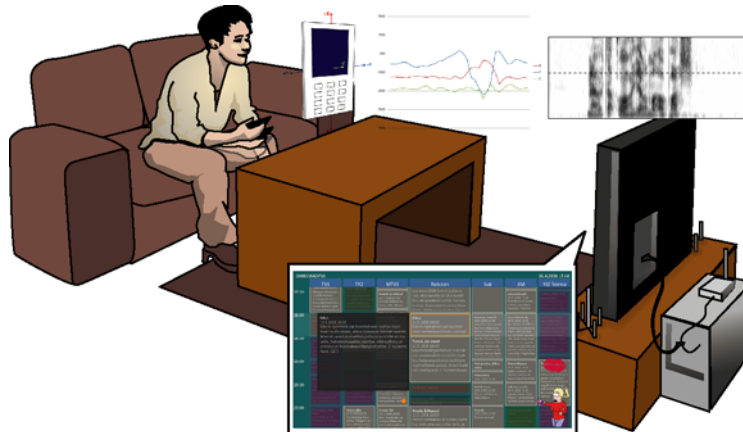


Fig. 1. Media center setup and the EPG user interface.

Users have full control over the media center via its speech user interface. For example, it is possible to record multiple episodes with a single utterance (“*Record all the Tom the Tractor shows this week*”), and to highlight programs based on their genre (“*Show me all the children programs tomorrow morning*”). The mobile phone keypad and gestures can be used for navigation and selections either independently or combined with each other. Gestures are made by moving the mobile phone in specific patterns. For visually impaired users, the system provides a tight integration between the Focus + Context GUI and synthesized speech output. Both the telephone and the mobile phone include speech output capabilities. In addition, haptic feedback is given using the vibration component of the mobile phone using a set of nine haptic patterns (hapticons). For physically impaired users, we have implemented a wheelchair mounted wireless microphone solution with voice activity and blow detection.

The first version of the media center application was demonstrated in a long-term public pilot study in a local museum (Rupriikki, Tampere) between June 2008 and March 2009. In addition, laboratory experiments have been carried out [2]. Here, we demonstrate a fully functional live system with additional functionality implemented for later pilot studies with physically and visually impaired users.

This work is supported by the Technology Development Agency of Finland (TEKES) under the Ubicom-programme in the "Ambient Intelligence Based on Sound, Speech and Multisensor Interaction"-project (TÄPLÄ, grant 40223/07).

References

1. Turunen, M., Hakulinen, J., Hella, J., Rajaniemi, J.-P., Melto, A., Mäkinen, E., Rantala, J., Heimonen, T., Laivo, T., Soronen, H., Hansen, M., Valkama, P., Miettinen, T., and Raisamo, R. Multimodal Media Center Interface based on Speech, Gestures and Haptic Feedback. In Proceedings of Interact 2009.
2. Turunen, M., Melto, A., Hella, J., Heimonen, T., Hakulinen, K., Mäkinen, E., Laivo, T., and Soronen, H. User Expectations and User Experience with Different Modalities in a Mobile Phone Controlled Home Entertainment System. In Proceedings of MobileHCI 2009.