Pr@senZ - P@CE: Mobile Interaction with Virtual Reality

Matthias Weber
International School of New Media (ISNM)
University of Lübeck, Germany
weber@isnm.de

Thies Pfeiffer
SFB 360
University of Bielefeld, Germany
Thies.Pfeiffer@Uni-Bielefeld.de

Bernhard Jung
International School of New Media (ISNM)
University of Lübeck, Germany
jung@isnm.de

ABSTRACT

Recently videoconferencing has been extended from human face-to-face communication to human machine interaction with Virtual Environments [2]. Relying on established videoconferencing protocol standards this thin client solution does not require specialized 3D soft- or hardware and scales well to multimedia enabled mobile devices. This brings a whole range of new applications to the mobile platform. To facilitate our research in mobile interaction the Open Source project P@CE has been started to bring a full-featured videoconferencing client to the Pocket PC platform.

Categories and Subject Descriptors

H.5.3 [**Information Interfaces**]: Computer-supported cooperative work; I.3.7 [**Computer Graphics**]: Virtual reality; H.4.3 [**Information Systems**]: Videoconferencing

General Terms

Algorithms, Performance

Keywords

Videoconferencing, shared virtual environments, collaboration, interaction, Pocket PC

1. INTRODUCTION

Enabling remote human computer interaction with Virtual Environments (VEs) fosters a broad range of new applications, e.g. for virtual help desks, edutainment or collaborative work. An approach supporting a broad range of clients has recently been proposed by Pfeiffer et al. [2]. This thin client solution relies on a videoconferencing (VC) gateway coupled to a shared VE, based on the H.323 standard VC protocol as provided by the free Open Source library OpenH323. This approach not only provides new viewing and interaction possibilities to VEs, it also makes VEs accessible anywhere where VC clients are available.

Copyright is held by the author/owner.
MobileHCI'05, September 19–22, 2005, Salzburg, Austria.
ACM 1-59593-089-2/05/0009.

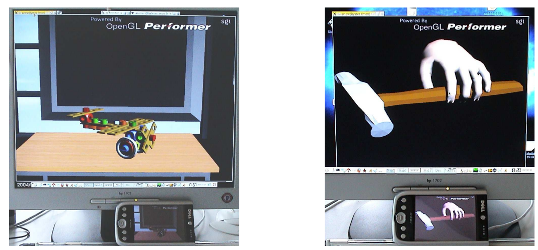

Figure 1: Viewing and controlling a collaborative Virtual Environment on a PDA

An example of a VC interface enabled application is the collaborative construction application shown in Fig. 1. The existing solution targeted at real-time render clusters and desktop PCs can now be accessed remotely and interacted with using typed text and speech recognition via the Internet from any H.323 VC software. With this approach basic audio and video processing capabilities are all that is needed on the client side, which makes this solution interesting for multimedia enabled mobile devices.

While VC software is already widespread on personal computers, the support for handheld devices is restricted to a few commercial applications or the abandoned PocketBone project, which does not support modern hardware properly. Other projects like Studierstube [3] use clients with a non-standard protocol to transfer content such as tracking data in addition to VC data. Using standard protocols, the Open Source project Pr@senZ/P@CE has been started to unfold the power of VC based interaction with VEs. The objective of Pr@senZ is to provide portable VC software using modern audio and video codecs and supporting different modalities, such as speech or text input. P@CE is the incarnation of Pr@senZ on the Pocket PC.

2. CONCEPTS AND REQUIREMENTS

Major requirements for our proposal of a VC interface to a VE, using a PDA, include: High performance graphics using modern VGA-capable displays and audio, video and textual input for providing several input modalities. Support for natural interaction, e.g. by way of speech and gesture input is also of high interest to us.

Bringing the VE to the User

For content output the remote participant should be able to access the VE at least by visual and auditory means.

To see and hear a VE on PDAs equipped with VC software the video and audio channels have to be connected to the VE. Pfeiffer et al. [2] already fulfilled this task by building a VC engine for VEs. This way a view into the VE gets transmitted via the video channel and sound events are played on the audio channel. On the PDA these video and audio channels can be seen and listened to via the VC system developed in this work. As a consequence the VE is brought to the user via standard VC technology.

Interaction with VEs

In addition to providing a means to transmit the VE to the remote user, VC functionalities can also be used to transmit user interactions to the VE on the server, as illustrated in the following examples.

Some VEs, such as [1], offer functionalities that allow the user to modify the scene content via natural language instructions. In such cases, the framework of Pfeiffer et al. [2] enables the remote user to verbally interact with the VE which is connected to a speech-recognition system. If the audio signal is too noisy and speech-recognition not appropriate, the participant can fall back on a textual chat protocol which is supported by most VC standards. In either way the incoming instruction can then be processed further by the VE.

Additionally, in the future the participant can interact visually with the VE, e.g. based on gesture input. While this seems quite demanding, a promising approach is applied in recent webcam based video games such as EyeToy [4]: here video signals of the participants are fed back into the content channel to allow the participants to synchronize their gestures with the content. Note that gesture processing does not require special support within the clients. Instead they rely on heavy image recognition on the server side.

3. REALIZATION

After an evaluation by Weber and Jung [5], OpenH323, a free implementation of the H.323 VC protocol, was chosen for Pr@senZ/P@CE. High performance video output in P@CE is realized using PocketHAL, which provides simplified, yet fast access to the frame buffer of various PDAs. Apart from this such a system shall remain platform-independent while providing low latencies and slender bandwidth consumption.

P@CE shall support several input modalities. They are recorded on the client side and after an appropriate, connection specific encoding, they are transmitted via VC channels to the VC engine on the VE server (see section 2). The recording part and how this data can be further processed in the VE will be described in the following.

Support for visual input modalities (e.g. gestures) necessitates the recording and transmission of video from the client device. Capturing video on PDAs is not a simple task as no standard interface exists. P@CE should at least support one type of camera that can be attached to a PDA. This way we can use the PDA to investigate new possibilities to control the VE. One application is to navigate in a VE just by moving the PDA. Another application is controlling the navigation in a VE through the user moving in front of the PDA. First steps have been taken to accomplish this task.

For audible modalities (speech) audio input has to be realized. This is supported by the underlying software architecture (Pocket PC and OpenH323) in a standardized way. Speech input can be transmitted to the server where a speech recognizer can convert the spoken words into commands to the VE. Of course, the PDA has to be equipped with a microphone so that this input modality can be used. If the recognition rate is poor the user may switch to textual input to give proper commands.

4. CONCLUSION

Interaction with Virtual Reality via videoconferencing is a promising technology. With a full featured videoconferencing client for the Pocket PC, this technology could be transferred to a highly mobile platform and enable novel types of applications. The Pr@senZ/P@CE initiative described in this paper aims to create such a videoconferencing solution for the Pocket PC. A first prototype is shown in Fig. 2. Pr@senZ/P@CE is available at http://pace.sourceforge.net.

Figure 2: A first live image from a Virtual Environment using the P@CE prototype

5. ACKNOWLEDGMENTS

This research is partially supported by the Deutsche Forschungsgemeinschaft (DFG) in the projects “Virtual Workers” and “Virtuelle Werkstatt” as well as the Collaborative Research Center SFB 360.

6. REFERENCES

[1] B. Jung, M. Hoffenke, and I. Wachsmuth. Virtual Assembly with Construction Kits. In *Proceedings of the 1998 ASME Design for Engineering Technical Conferences (DECT-DFM ’98)*, 1998.

[2] T. Pfeiffer, M. Weber, and B. Jung. Ubiquitous Virtual Reality: Accessing Shared Virtual Environments through Videoconferencing Technology. In M. McDerby and L. Lever, editors, *Proceedings of the EG UK Theory and Practice of Computer Graphics*. Eurographics, 2005.

[3] D. Schmalstieg, A. Fuhrmann, G. Hesina, Z. Szalavari, L. M. Encarnação, M. Gervautz, and W. Purgathofer. The Studierstube Augmented Reality Project. In *PRESENCE - Teleoperators and Virtual Environments*, volume 11. MIT Press, 2002.

[4] Sony Computer Entertainment Inc. EyeToy. Web address: http://www.eyetoy.com, 2005.

[5] M. Weber and B. Jung. Implementierung kollaborativer Mixed-Reality-Systeme: Ein Vergleich von Open-Source Videokonferenz-Software. In *Virtuelle und Erweiterte Realität - 1. Workshop der GI-Fachgruppe VR/AR*, pages 173–180, Aachen, 2004. Shaker-Verlag.