



biblio.ugent.be

The UGent Institutional Repository is the electronic archiving and dissemination platform for all UGent research publications. Ghent University has implemented a mandate stipulating that all academic publications of UGent researchers should be deposited and archived in this repository. Except for items where current copyright restrictions apply, these papers are available in Open Access.

This item is the archived peer-reviewed author-version of:

Immersive Video Conferencing Architecture using Game Engine Technology

Chris Poppe, Charles-Frederik Hollemeersch, Sarah De Bruyne, Peter Lambert, and Rik Van de Walle

In: *Lecture Notes in Computer Science*, 6523(2011), 486-488, 2011.

Optional: <http://www.springerlink.com/content/mr77k50k50363287/>

To refer to or to cite this work, please use the citation to the published version:

Chris Poppe, Charles-Frederik Hollemeersch, Sarah De Bruyne, Peter Lambert, and Rik Van de Walle (2011). Immersive Video Conferencing Architecture using Game Engine Technology. *Lecture Notes in Computer Science* 6523(2011) 486-488. DOI: 10.1007/978-3-642-17829-0_48

Immersive Video Conferencing Architecture using Game Engine Technology

Chris Poppe, Charles-Frederik Hollemeersch, Sarah De Bruyne, Peter Lambert,
and Rik Van de Walle

Multimedia Lab, Ghent University - IBBT,
Gaston Crommenlaan 8, B-9050 Ledeborg-Ghent, Belgium
{chris.poppe, charlesfrederik.hollemeersch, sarah.debruyne, peter.lambert,
rik.vandewalle}@ugent.be
<http://multimedialab.elis.ugent.be>

Abstract. This paper introduces the use of gaming technology for the creation of immersive video conferencing systems. The system integrates virtual meeting rooms with avatars and live video feeds, shared across different clients. Video analysis is used to create a sense of immersiveness by introducing aspects of the real world in the virtual environment. This architecture will ease and stimulate the development of immersive and intelligent telepresence systems.

Keywords: Video Conferencing, Game Engine, Video Analysis

1 Introduction

Telepresence allows a person to feel as if they were present at a location other than their true location. Current systems lack in conveying the true telepresence feeling since they just visualize the other meeting rooms. Recent work shows the interest in more immersive telepresence by creating virtual worlds or 3D representations of participants [1]. Additionally, efforts have been done to introduce different information sources to video conferencing. Typical examples are the application of video analysis for head or gaze tracking [2, 3].

In this work we show that a game engine is well-suited to create such immersive telepresence systems and allows for easy updating, modular architectures, advanced networking and rendering capabilities. The work is part of the iCocoon (immersive Communication through Computer vision) project ¹.

2 Video Conferencing using Game Engine Technology

Current game engines allow the development of 3D multi-player games. They provide means for creating, editing and rendering virtual worlds, interaction and network exchange of information. Created for the ease of game development, we

¹ <http://www.ibbt.be/en/projects/overview-projects/p/detail/icocoon-2>

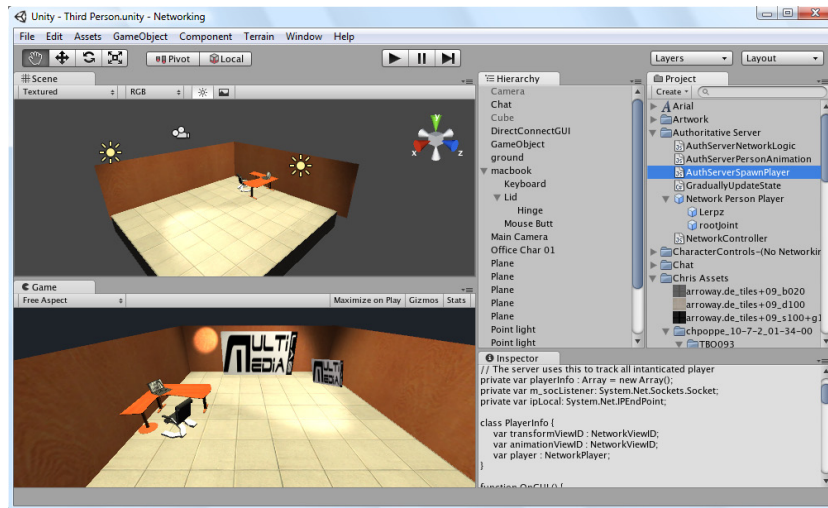


Fig. 1. Application of Unity for immersive video conferencing

propose to benefit from the efforts made in this area to create an immersive video conferencing system. For this purpose, *Unity*², a multi-player game development tool, is used to create game applications (see Fig. 1 for an example).

Each of those applications can act as server or client and present the end-user with a view on a virtual environment consisting of avatars, life video feeds, and models of the meeting rooms. The user can interact with the virtual world on different levels (e.g., chatting, logging in and off, and selecting different camera views). The latter can be accomplished using C#-script, inherent to Unity to implement game functionality. Both the virtual world and avatars are synchronized using the built-in network support of Unity (remote procedure calls and state synchronization), allowing to shield the developer from low-level network programming. The video feeds are also analysed by external video analysis modules which extract relevant information, (using OpenCV³). Different analysis scenarios can be imagined, e.g., detection of illumination, motion of the camera, moving objects, face detection, gaze tracking, and gesture recognition. The results of the analysis is represented as metadata and send to the server over socket communication. The server consequently processes the metadata and makes appropriate changes to the virtual world. Using the game technology, these changes are automatically communicated to the other clients. Changes of the avatars, meeting rooms, and camera viewpoints are typical examples of actions that are triggered by the metadata.

Figure 2 shows screen shots of a possible set-up. A virtual meeting room is rendered with avatars and life video feeds for each connected client. When the

² <http://unity3d.com/unity/>

³ <http://opencv.willowgarage.com/wiki/>

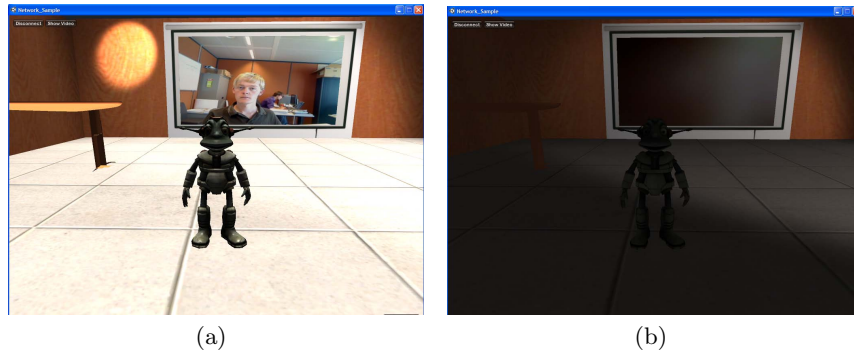


Fig. 2. Screen shot of the virtual meeting room with video feed.

light is turned off in the real world, the illumination change is detected by video analysis and send to the server. The server updates the light source in the virtual environment, which is consequently reflected on all clients.

3 Conclusions

We propose the use of game development tools to create immersive video conferencing. This allows to create virtual meeting rooms with avatars shared across different clients. Additionally, video analysis is integrated to create a sense of immersiveness by introducing aspects of the real world in the virtual environment. Future work consists of testing the system with multiple users on aspects like scalability, performance, and quality of user experience.

Acknowledgments The research activities that have been described in this paper were funded by Ghent University, the Interdisciplinary Institute for Broadband Technology (IBBT), the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT-Flanders), the Fund for Scientific Research-Flanders (FWOFlanders), and the European Union.

References

1. Schreer, O., Feldmann, I., Atzpadin, N., Eisert, P., Kauff, P., Belt, H.J.W.: 3D presence - a system concept for multi-user and multi-party immersive 3D video conferencing. IET 5th European Conference on Visual Media Production, pp. 10–17 (2008)
2. Kompatsiaris, I., Strintzis, M.G.: Spatiotemporal Segmentation and Tracking of Objects for Visualization of Videoconference Image Sequences. IEEE Trans. on Circuits and Systems for Video Technology. 10, pp. 1388–1402 (2000)
3. Vertegaal, R., Weevers, I., Sohn, C., Cheung, C.:GAZE-2: Conveying Eye Contact in Group Video Conferencing Using Eye-Controlled Camera Direction. In: Conference on Human Factors in Computing Systems, pp. 521 – 528. ACM (2003)