A SPATIAL IMMERSIVE OFFICE ENVIRONMENT FOR COMPUTER-SUPPORTED COLLABORATIVE WORK

Moving Towards the Office of the Future

Maarten Dumont[1], Sammy Rogmans[1,2], Steven Maesen[1], Karel Frederix[1],
Johannes Taelman[1] and Philippe Bekaert[1]
[1]*Hasselt University, tUL, IBBT, Expertise Centre for Digital Media, Wetenschapspark 2, 3590 Diepenbeek, Belgium*
[2]*Multimedia Group, IMEC, Kapeldreef 75, 3001 Leuven, Belgium*

Keywords: Immersive environment, Office of the future, Multitouch surface, Virtual camera, Multiprojector.

Abstract: In this paper, we present our work in building a prototype office environment for computer-supported collaborative work, that spatially – and auditorially – immerses the participants, as if the augmented and virtual generated environment was a true extension of the physical office. To realize this, we have integrated various hardware, computer vision and graphics technologies from either existing state-of-the-art, but mostly from knowledge and expertise in our research center. The fundamental components of such an office of the future, i.e. image-based modeling, rendering and spatial immersiveness, are illustrated together with surface computing and advanced audio processing, to go even beyond the original concept.

1 INTRODUCTION

In these modern times, professional collaboration between people becomes more and more a necessity, very often even over long distances. This creates the need for a futuristic office environment where people can instantly collaborate on demand, as if they were in the same time and place, anywhere and anytime (see Fig. 1). The visions of an office of the future are not new, and most famous from the well-known publication of(Raskar et al., 1998), envisioning the unification of computer vision and graphics. A general consensus in this research domain is that there are 3 fundamental requisites to crystalize these ideas:

- **Dynamic Image-based Modeling.** The computer vision part that analyzes the scenery and potential display surfaces in the related offices and dynamically models them in real-time for further processing or augmentation with virtual images.
- **Rendering.** The computer graphics part that renders the models according to the analyzed (irregular) display surfaces, position and viewing direction of the participants.
- **Spatially Immersive Display.** The hardware part that provides with a sufficient immersive medium for the virtual rendered or augmented images of the generated models and physical scenery.

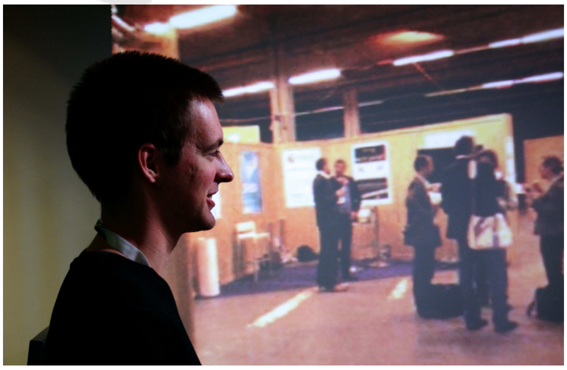

Figure 1: Teaser picture of being immersed within a large real-time virtual office environment.

Although the ideas themselves are not new, they remain very challenging because it requires the integration of advanced computer vision, graphics and hardware, moreover, all functioning seemlessly together within a real-time constraint.

Dumont M., Rogmans S., Maesen S., Frederix K., Taelman J. and Bekaert P..
A SPATIAL IMMERSIVE OFFICE ENVIRONMENT FOR COMPUTER-SUPPORTED COLLABORATIVE WORK - Moving Towards the Office of the Future.
DOI: 10.5220/0003567702120216
In *Proceedings of the International Conference on Signal Processing and Multimedia Applications* (SIGMAP-2011), pages 212-216
ISBN: 978-989-8425-72-0

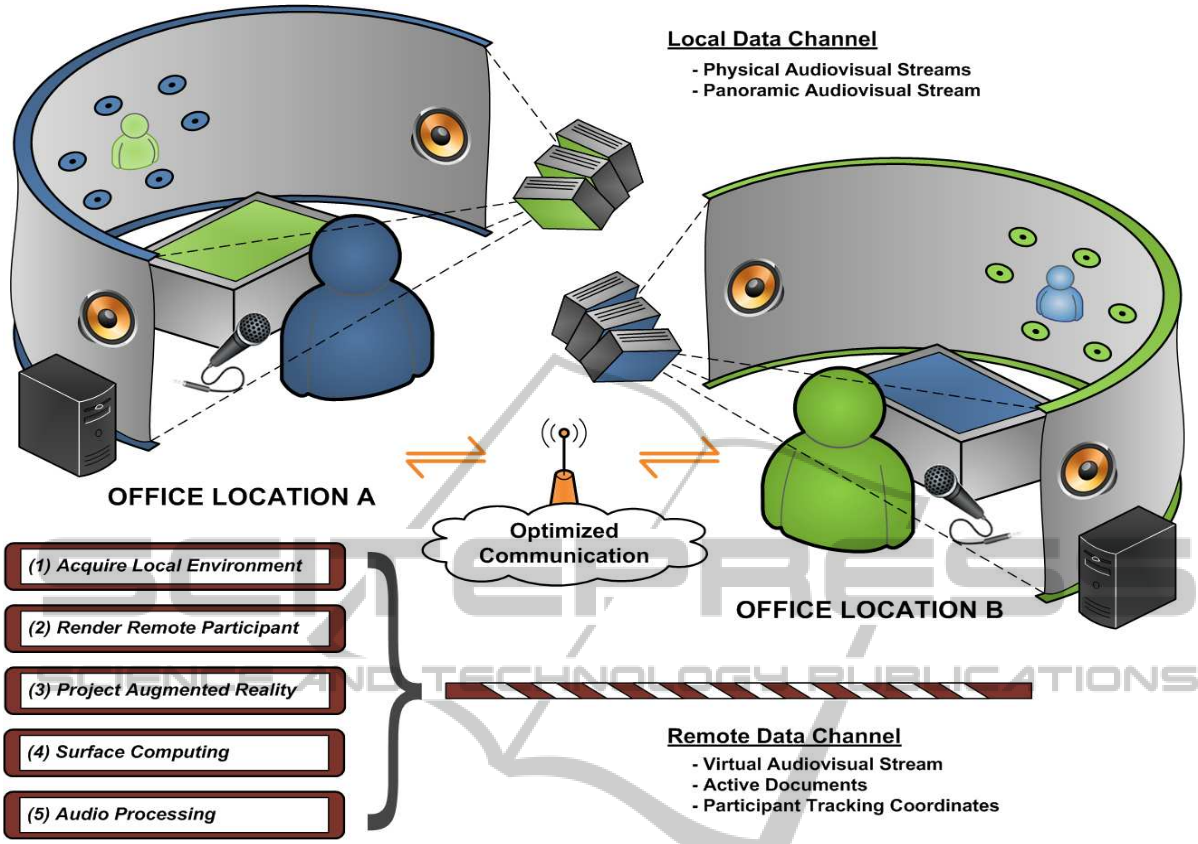

Figure 2: The prototype of our office environment, existing out of 5 application building blocks: (1) Acquire local environment, (2) render remote participant, (3) project augmented reality, (4) surface computing, and (5) audio processing.

2 OFFICE ENVIRONMENT

We have developed a prototype (see Fig. 2) of a spatial immersive office environment that is coming close to computer-supported collaborative work like it has been long envisioned in the conceptual office of the future (Raskar et al., 1998). We implement the three fundamental requisites of these futuristic working environments, using a sea of cameras that lies at the base of our reconstruction and rendering algorithm that (1) dynamically models and (2) renders the required scenery at once. Furthermore, we have created a spatial immersive display using a multiprojector setup that (3) projects an augmented reality to form a virtual extension of the physical office. Going beyond the original concept, we integrated (4) surface computing and (5) audio processing for realistic communication.

2.1 Acquire Local Environment

As the first requisite of futuristic immersive offices, we place a sea of cameras to acquire and dynamically model the local environment. We constrain the panoramic display surface as being fixed, which is often the case in practical situations, and therefore do not need a constant autocalibration of the projectors. Hence the sea of cameras can be limited to only 4 or 6 pieces without impeding on the resulting quality of the generated virtual imagery, i.e. modelling and rendering the office participant. For minimal interference with the environment, the cameras are placed behind the panoramic display surface, while small holes in the screen allow the lenses to slightly pale through and capture the physical scenery.

Instead of contructing a genuine 3D model of the participant, we exploit the determined coordinates of a high-speed person and eye tracking module to only synthesize the required point of view for the remote participant. A drastic amount of modelling and rendering computations are therefore bypassed, moreover, the physical (audio)visual streams are used locally, and only the virtual (audio)visual stream is sent over the network. Hence, the data processing and communication is optimized to guarantee the real-time aspect of the office system, even when using inexpensive commodity processing hardware.

2.2 Render Remote Participant

As the second requisite of immersive office collaboration, the remote participant is rendered correctly according to the position and viewing location of the local user. We do this by using an intelligent and optimized plane sweeping algorithm (Yang et al., 2002; Dumont et al., 2008; Dumont et al., 2009b; Rogmans et al., 2009b) that harnesses the computational power of the massive parallel processing cores inside contemporary graphics cards (Owens et al., 2008; Rogmans et al., 2009c; Rogmans et al., 2009a; Goorts et al., 2009; Goorts et al., 2010) for maintaining the real-time constraint. Furthermore, the rendering is made so that the eye contact between the collaborators is restored, without physically having to look inside the camera lens. The immersivity is therefore already quite high, and can be further improved by optionally rendering the remote participant stereoscopically for natural 3D perception (Dumont et al., 2009a; Rogmans et al., 2010a). However, as this requires the inconvience of wearing active shutter glasses, we often resort to exploiting only monocular depth cues while still perceiving good 3D (Held et al., 2010; Rogmans et al., 2010b).

Both participants are augmented with an ad hoc precaptured panoramic environment to truly immerse the users. Nonetheless, the (audio)visual stream can be kept locally as the virtual context of the stream serves the sole purpose of consistently expanding the given physical office space. The high-speed person and eye tracking can therefore also take care of the dynamic rendering, following mainly the rules of motion paralax, yet optionally also other important natural depth cues to further maximize the credibility of the virtual environment background being real.

2.3 Project Augmented Reality

The third fundamental requisite of a futuristic office is a spatial immersive display, such as e.g. CAVE (Cruz-Neira et al., 1993; Juarez et al., 2010), Cavelet (Green and Whites, 2000) or Blue-C (Gross et al., 2003). We managed to build a rather cheap immersive display by spanning a durable heavy white vinyl cloth over a series of lightweight aluminum pipes, to form a 180 degree panoramic screen using multiple projectors. To supress the overall cost, the cloth is matte white with reflection less than 5%, instead of the cinematic pearlescent screens with a reflection of about 15%. This results in the fact that black is observed as a form of dark grey and the overall brightness as rather low. As a consequence, we do not rely on subliminal imperceivable structured light that is embedded in the projector feeds. Nevertheless, we do use some additional flood lights that are carfully placed above the panoramic screen, to ensure proper lighting of the office participants.

The lack of constant (imperceivable) structured light renders it impossible to continously autocalibrate the projectors for dynamic image correcting if the cloth should change shape. However, as the cloth is spanned tightly around the aluminum pipes, the panoramic display is fixed so that the projectors only need to be calibrated once at the office setup. The multiprojector setup is calibrated to be able to seemlessly stitch the multiple projections and to geometrically correct the image according to the shape of the screen. While various methods already exist (Raskar et al., 1999; Fiala, 2005; Harville et al., 2006; Griesser and Gool, 2006; Sajadi and Majumder, 2010), our single ad hoc calibration is based on perceivable structured light that is being recorded by a temporary camera, which is placed within the vicinity of the projectors.

2.4 Surface Computing

Beyond the three fundamental requisites to build an office of the future, we also integrated cooperative surface computing (Dietz and Leigh, 2001; Cuypers et al., 2009; Wobbrock et al., 2009). In contrast to standard outdated computer-supported collaborative work systems that typically only share software applications, the networked surface computing allows the office participants to genuinely exchange documents as if they were interacting with printed paper. The used surface computer features a typical multitouch control, commonly known on devices such as the Apple iPhone, iPod touch and iPad, providing a natural and intuitive feel when handling documents. The files that are opened are shared between the users by the network and can be viewed, controlled, manipulated and annotated simultaneously for true immersive collaboration, as if the particpants were working at the same table. Although it is not a direct criterion in the original draft of the office of the future, it greatly contributes to the collaborative characteristics and is an essential part of futuristic office environments.

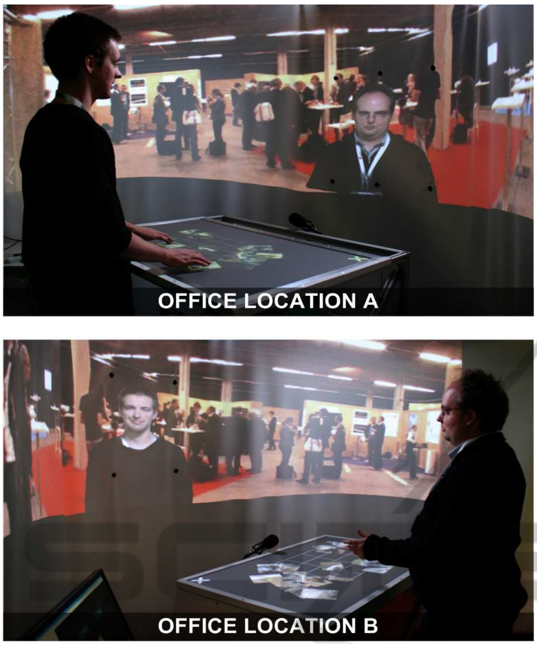

Figure 3: Picture of our prototype setup when collaborating between office location A and B.

2.5 Audio Processing

As a final part of our futuristic office prototype, we included advanced audio processing to further facilitate and complete the communication between the users. For now, we only have support for monaural sound, as our office is designed for one-to-one collaboration. The sound is hence captured with a single high-fidelity microphone and send to the other side for processing. Upon arrival, the audio stream is first registered and synchronized before being amplified and outputted through the speakers. As the sound is played, the Larsen effect, i.e. the audio feedback that is generated through the loop created between both audio systems, is dynamically cancelled using the determined network transfer delay at registration. The synchronization and echo cancellation contributes in a natural way of communicating, giving the users the feeling of talking to each other in the same room.

While not yet implemented, the system design lends itself perfectly for genuine 3D audio by using a minimum of 2 microphones and reconstructing the 3D sound by according surround speaker setups. This also provides the office participants the direction of the speaker, which is particulary useful at many-to-many collaborations. However, the sound must remain consistent with the augmented reality.

3 PROTOTYPE RESULTS

Our system was initially built and tested at our research lab, but was also demonstrated and used at the ServiceWave convention in December 2010 (see Fig. 3). Per office setup we used two computers, one for the acquisition and rendering, and one for the immersive projection. Both computers had an Intel Core 2 Quad CPU and a GTX280 graphics card of NVIDIA with 1GB GDDR3 memory. The cameras were Point Grey Grashoppers and Flees with a FireWire connection, and the projectors Optoma TX1080s. The surface computer had an embedded PC for displaying the documents and processing the mutitouch gestures, while the audio processing and echo cancellation was done on an individual Mac Mini. In practice, our system achieved real-time speeds over 26 fps.

A demostration movie of our futuristic office environment, as presented at ServiceWave 2010, can be found on http://research.edm.uhasselt.be/∼mdumont/Sigmap2011. We invite the reader to have a look in order to get a better understanding of the possibilities and complexity of the system.

4 CONCLUSIONS

We have presented our prototype of a futuristic office for computer-supported colalborative work in a contraint environment. Even though we implemented the fundamental criteria as originally stated by (Raskar et al., 1998), we even went beyond by additionally using surface computing and advanced audio processing, while still achieving over 26 fps real-time speed.

ACKNOWLEDGEMENTS

We would like to acknowledge the Interdisciplinary Institute for Broadband Technology (IBBT). Co-author Sammy Rogmans is funded by a specialization bursary from the IWT, under grant number SB073150 at the Hasselt University.

REFERENCES

Cruz-Neira, C., Sandin, D., and DeFanti, T. (1993). Virtual reality: The design and implementation of the cave. In *ACM SIGGRAPH 93*, pages 135–142.

Cuypers, T., Frederix, K., Raymaekers, C., and Bekaert, P. (2009). A framework for networked interactive surfaces. In *2nd Workshop on Software Engineering and Architecture for Realtime Interactive Systems (SEARIS@VR2009)*.

Dietz, P. and Leigh, D. (2001). Diamondtouch: A multi-user touch technology. In *14th Annual ACM Symposium on User Interface Software and Technology*, Florida, FL, USA.

Dumont, M., Maesen, S., Rogmans, S., and Bekaert, P. (2008). A prototype for practical eye-gaze corrected video chat on graphics hardware. In *International Conference on Signal Processing and Multimedia Applications*, Porto, Portugal.

Dumont, M., Rogmans, S., Lafruit, G., and Bekaert, P. (2009a). Immersive teleconferencing with natural 3d stereoscopic eye contact using gpu computing. In *3D Stereo Media*, Liege, Belgium.

Dumont, M., Rogmans, S., Maesen, S., and Bekaert, P. (2009b). Optimized two-party video chat with restored eye contact using graphics hardware. *CCIS*, 48:358–372.

Fiala, M. (2005). Automatic projector calibration using self-identifying patterns. *Computer Vision and Pattern Recognition Workshop*, 0:113.

Goorts, P., Rogmans, S., and Bekaert, P. (2009). Optimal data distribution for versatile finite impulse response filtering on next-generation graphics hardware using cuda. In *The Fifteenth International Conference on Parallel and Distributed Systems*, pages 300–307, Shenzhen, China.

Goorts, P., Rogmans, S., Eynde, S. V., and Bekaert, P. (2010). Practical examples of gpu computing optimization principles. In *International Conference on Signal Processing and Multimedia Applications*, Athens, Greece.

Green, M. and Whites, L. (2000). The cave-let: a low-cost projective immersive display. *Journal of Telemedicine and Telecare*, 6(2):24–26.

Griesser, A. and Gool, L. V. (2006). Automatic interactive calibration of multi-projector-camera systems. *Computer Vision and Pattern Recognition Workshop*, 0:8.

Gross, M., Würmlin, S., Naef, M., Lamboray, E., Spagno, C., Kunz, A., Koller-Meier, E., Svoboda, T., Van Gool, L., Lang, S., Strehlke, K., Moere, A. V., and Staadt, O. (2003). blue-c: a spatially immersive display and 3d video portal for telepresence. *ACM Trans. Graph.*, 22:819–827.

Harville, M., Culbertson, B., Sobel, I., Gelb, D., Fitzhugh, A., and Tanguay, D. (2006). Practical methods for geometric and photometric correction of tiled projector. In *2006 Conference on Computer Vision and Pattern Recognition Workshop*, CVPRW '06, pages 5–, Washington, DC, USA. IEEE Computer Society.

Held, R. T., Cooper, E. A., O'Brien, J. F., and Banks, M. S. (2010). Using blur to affect perceived distance and size. *ACM Trans. Graph.*, 29:19:1–19:16.

Juarez, A., Schonenberg, B., and Bartneck, C. (2010). Implementing a low-cost cave system using the cryengine2. *Entertainment Computing*, 1(3–4):157–164.

Owens, J. D., Houston, M., Luebke, D., Green, S., Stone, J. E., and Phillips, J. C. (2008). Gpu computing. *Proceedings of the IEEE*, 96(5):879–899.

Raskar, R., Brown, M. S., Yang, R., Chen, W.-C., Welch, G., Towles, H., Seales, B., and Fuchs, H. (1999). Multi-projector displays using camera-based registration. In *Proceedings of 10th IEEE Visualization 1999 Conference (VIS '99)*, VISUALIZATION '99, Washington, DC, USA. IEEE Computer Society.

Raskar, R., Welch, G., Cutts, M., Lake, A., Stesin, L., and Fuchs, H. (1998). The office of the future: a unified approach to image-based modeling and spatially immersive displays. In *Proceedings of 25th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '98, pages 179–188, New York, NY, USA. ACM.

Rogmans, S., Bekaert, P., and Lafruit, G. (2009a). A high-level kernel transformation rule set for efficient caching on graphics hardware - increasing streaming execution performance with minimal design effort. In *International Conference on Signal Processing and Multimedia Applications*, Milan, Italy.

Rogmans, S., Dumont, M., Cuypers, T., Lafruit, G., and Bekaert, P. (2009b). Complexity reduction of real-time depth scanning on graphics hardware. In *VISAPP*, pages 547–550, Lisbon, Portugal.

Rogmans, S., Dumont, M., Lafruit, G., and Bekaert, P. (2009c). Migrating real-time image-based rendering from traditional to next-gen gpgpu. In *3DTV-CON: The True Vision Capture, Transmission and Display of 3D Video*, Potsdam, Germany.

Rogmans, S., Dumont, M., Lafruit, G., and Bekaert, P. (2010a). Biological-aware stereoscopic rendering in free viewpoint technology using gpu computing. In *3DTV-CON: The True Vision Capture, Transmission and Display of 3D Video*, Tampere, Finland.

Rogmans, S., Dumont, M., Lafruit, G., and Bekaert, P. (2010b). Immersive gpu-driven biological adaptive stereoscopic rendering. In *3D Stereo Media*, Liege, Belgium.

Sajadi, B. and Majumder, A. (2010). Auto-calibration of cylindrical multi-projector systems. In *Virtual Reality Conference*, Waltham, MA, USA.

Wobbrock, J. O., Morris, M. R., and Wilson, A. D. (2009). User-defined gestures for surface computing. In *27th international Conference on Human Factors in Computing Systems*, New York, NY, USA.

Yang, R., Welch, G., and Bishop, G. (2002). Real-time consensus-based scene reconstruction using commodity graphics hardware. In *PG '02: Proceedings of the 10th Pacific Conference on Computer Graphics and Applications*, page 225, Washington, DC, USA. IEEE Computer Society.