IMPROVE: Collaborative Design Review in Mobile Mixed Reality

Pedro Santos[1], André Stork[1], Thomas Gierlinger[2], Alain Pagani[2], Bruno Araújo[3], Ricardo Jota[3], Luis Bruno[3], Joaquim Jorge[3], Joao Madeiras Pereira[3], Martin Witzel[4], Giuseppe Conti[4], Raffaele de Amicis[4], Iñigo Barandarian[5], Céline Paloc[5], Oliver Machui[6], Jose M. Jiménez[7], Georg Bodammer[8], and Don McIntyre[9]

[1] Fraunhofer-IGD, A2
[2] TU-Darmstadt, FB21,GRIS
[3] INESC-ID
[4] GraphiTech
[5] VICOMTech.,
[6] Trivisio GmbH
[7] MED
[8] STT,
[9] Lighthouse
{Pedro.Santos, Andre.Stork}@igd.fhg.de, {Thomas.Gierlinger, Alain.Pagani}@igd.fhg.de, {brar, jota.costa, bruno, jaj, jap}@inesc-id.pt, {martin.witzel, giuseppe.conti, raffaele.de.amicis}@graphitech.it, {Inigo.Barandarian, Celine.Paloc}@vicomtech.es, machui@trivisio.de, georg.bodammer@microemissive.com, jmjimenez@stt.es, don.mcintyre@urbanlearningspace.com

Abstract. In this paper we introduce an innovative application designed to make collaborative design review in the architectural and automotive domain more effective. For this purpose we present a system architecture which combines variety of visualization displays such as high resolution multi-tile displays, TabletPCs and head-mounted displays with innovative 2D and 3D Interaction Paradigms to better support collaborative mobile mixed reality design reviews. Our research and development is motivated by two use scenarios: automotive and architectural design review involving real users from Page\Park architects and FIAT Elasis. Our activities are supported by the EU IST project IMPROVE aimed at developing advanced display techniques, fostering activities in the areas of: optical see-through HMD development using unique OLED technology, marker-less optical tracking, mixed reality rendering, image calibration for large tiled displays, collaborative tablet-based and projection wall oriented interaction and stereoscopic video streaming for mobile users. The paper gives an overview of the hardware and software developments within IMPROVE and concludes with results from first user tests.

1 Introduction

Design Review is one of the most prominent areas benefiting from Virtual Reality and Immersive Projection Technologies. Use cases generally comprise many observers in

front of a back projection wall discussing a design decision on a virtual model. The user's visual experience is further enhanced by the possibility of individually rendering the 3D stereoscopic images of a model for each user's point of view. Recent developments allow up to 4 tracked users [Fröhlich04].

In the last couple of years Augmented Reality ceased to be used only for maintenance and training but also in design and design review (e.g. see ARVIKA project homepage: www.arvika.de). Typically video-based AR is applied to augment physical models with design alternatives e.g. new virtual dash boards (VW), alternative wheel designs (Audi). However Interaction and collaboration seldom take place in such applications.

SpaceDesign [Fiorentino02] was one of the first creative and generative tools allowing a user equipped with stereoscopic optical see-through glasses to sketch free-form surfaces directly in 3D augmenting a physical model underneath.

Schmalstieg [Schmalstieg 06] explored the possibilities of mobile collaborative AR within their Studierstube system supporting various applications from scientific visualisation to interactively experiencing math and esp. geometry.

With AR technology maturing, companies such as BMW [Klinker02] became curios and interested whether AR can be used in large(r) environments enabling mobile users to compare virtual with real models by walking around them in presentation facilities.

2   Application Scenarios

For us this was the starting point for brainstorming the possibilities of Mixed Reality in the design process with representatives from automotive industry and architecture. One thing important to note is that the two branches are completely different in structure. Car makers are big companies that to a great extent have pushed VR technology in the last 15 years whereas the architecture branch is characterized by many small enterprises that cannot afford expensive VR installations.

Discussions with architects revealed the following needs:

1. Starting a new Building: Early shape studies integrating new building in its surroundings and the landscape – Support by sketching capabilities on the construction site using Tablet-PCs for early sketches and optical see-through glasses for visualising the sketch in its physical environment.
2. Refinement of early sketches in the office: Collaborative Review and interactive presentation to customers using HMDs supporting multiple viewports. Pen-based direct 3D interaction is envisaged for creating annotations and change orders.
3. Final Design: Presentation of final design on construction site for multiple users. Again, HMDs and large area tracking is needed. The correct lighting of the virtual model with respect to the lighting conditions at the construction site is important.


Fig. 1. Onsite sketching and collaborative indoor reviewing scenarios (conceptual sketches)

In contrast to the architects, whose scenarios are centred around the use of HMDs indoor or on location, the representatives from automotive industry are familiar with large area displays and want to improve and extend the use of large single- or multi-tile projection technology.

The automotive scenarios look as follows:

1. The reviewers in front of the large projection wall shall be equipped with Tablet PCs. The Tablet PC shows the scene on the wall in reduced resolution from the same viewpoint unless a user decouples from the wall. Decoupled users can create annotations and suggestions for changes using handwritten input on their Tablet PC and "re-connect to the wall" at will.
2. Multiple users can work side-by-side to the wall in an AR setup performing direct 3D interaction with 3D pen.
3. Single users in front of the wall can directly interact with the wall using hand gestures.

To improve geometric and colour consistency of high resolution walls more efficient calibration techniques are required.
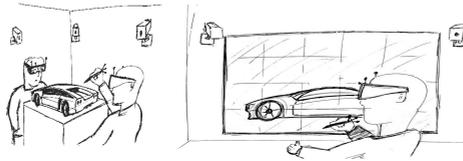

Fig. 2. Collaborative cAR design review scenario combining large and head mounted displays (conceptual sketches)

3 Requirements

Analysing the mobile collaborative scenarios (indoors and outdoors) from both branches entails that a combination of state-of-the art and progress in the following areas is needed to answer the requirements:

- lightweight, power efficient stereoscopic optical see-through HMDs
- large area multi user tracking
- augmented reality rendering
- collaborative interaction techniques spanning display walls, Tablet PCs and direct 3D interaction
- image calibration techniques
- image (video) transmission techniques to ensure highest possible rendering quality for mobile users using limited computational and graphics power

The authors are aware of no comparable approach that offers that unique combination of technologies and techniques. The most comparable work is probably done by Regenbrecht et al [Regenbrecht02] while IMPROVE is wider in scope and spans more diverse environments and scenarios. IMPROVE not only contributes to the mixed and augmented reality field with that combination of techniques listed

above but also with new achievements such as taylor-made organic light emitting display (OLEDs) by the consortium partner Micro Emissive Displays (MED), efficient image calibration techniques, innovative interaction techniques, pre-computed radiance based rendering and optimized video transmission for mobile users.

4   IMPROVE System Architecture

This chapter presents the general IMPROVE architecture (Fig. 3). We will analyze the different components subsequently.
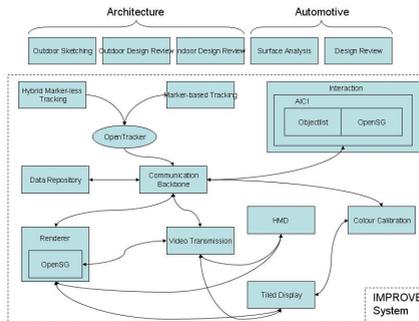

Fig. 3. IMPROVE System Architecture

Communication Backbone: The communication backbone is one of the internal modules which is part of the IMPROVE system i.e. the AICI framework and is responsible for providing to the other internal modules a simple interface to support the communication. This communication interface allows other modules or the main interface view of the application access to the following functionalities:

- Connection to the distributed server i.e. an instance of the XMLBlaster server available in the network
- Login and Logout to the distributed server using a username for identification
- Subscription of an interest to receive a network message of a given type
- Sending mechanism to publish messages of a given type to the network

Fig. 4. Creation of annotations and material property changes

Messages can be object updates, navigation messages or session messages, i.e.: one instance could broadcast an update to an already existing annotation on the OSGA communication backbone.

In the AICI system the communication is accessible trough the Communicator which is the internal module of the system that interacts with the OSGA communication layer which provides a simple interface to create an XMLBlaster Client with the ability to communicate with the distributed server which is an XMLBlaster server instance running in the network.

Interaction: The architectural indoor interaction subsystem supports interaction between users during collaborative reviews. The system operates as two or more IMPROVE instances connected, through OSGA, to the communication backbone. The IMPROVE instances interact by sending and receiving IMPROVE entity messages.

The scenario presents two or more architects in a collaborative design session or an architect and a client on a collaborative review session. In both situations (collaborative sessions), all actors use Tablet PCs. The Space Mouse is also supported in addition to the Tablet PC pen. Furthermore the scenario supports users with HMDs who visualize models on top of a table or outdoors and can be shown design alternatives by architects using TabletPCs. In the automotive scenario the setup may involve a collaborative design session in front of a powerwall. In all cases there are a variety of interaction modalities to provide input to the TabletPCs or powerwall.
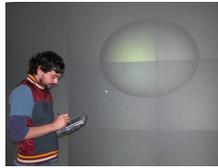
Fig. 5. Creation shapes and annotation in front of a Power-wall with a TabletPC

Collaborative entity messages can be divided into two groups: Shape editing and Annotation editing. Shape messages are only used during collaborative design sessions between architects. Clients are not allowed to create new shapes. Annotation entity messages are available on both design sessions and review sessions.

In the moment possible shapes are restricted to primitive shapes like cones, cylinders and cubes (Fig. 5). We plan to also be able to create free-form surfaces.

Annotations use the Post-It metaphor. Users can take notes of their comments and then stick them to objects. This is done by choosing an appropriate anchor on the object to stick the note to. Post-It notes will always face the user. All commands are broadcast to the OSGA communication backbone.

The IMPROVE system includes the following interaction modalities:

- Laser Interaction – For Powerwall interaction (Fig.6). TabletPC uses the pen for the same effect;
- Voice Activation commands – Available on both indoor and outdoor scenarios;
- Smart Widget – The Powerwall scenario relies on a mobile device input to create annotations. The mobile device is chosen over other solutions because of its lightweight, user familiarity and ubiquity.

- Multi-modal interaction developed focus on using Laser interaction as main interaction modality and Voice or Smart Widgets as backup modalities. The interaction techniques described in the following sections describe the solutions we implemented using the presented methods.

Annotation Interaction. Annotation interaction uses Laser Interaction as a pointing device and Smart Widgets as annotation input device.

Menu Interaction. Some actions will only be available though a menu. Unfortunately the standard way to access a menu - mouse right-click - is not available on the Powerwall scenario. To overcome this problem we devised a menu activation gesture to be executed with the laser - This interaction was also adapted to the Tablet PC pen interaction. To activate menu options a gate mechanism adopted. This allowed the users to navigate the menu and activate menu options on a single stroke, as depicted on Figure 6.
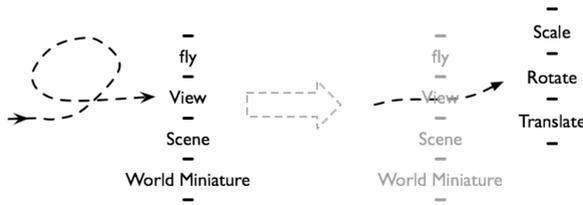
Fig. 6. Menu interaction example. User executes a gesture and a gate menu appears. New menu options appear in front on the laser postion for a fluid interaction.

Object Selection. To solve the laser limitation, we developed a novel selection interaction: To select an object you circle the object. In order to select several objects, just selected one or more at the same time. Moreover, the selections are cumulative, thus each selection adds the objects to the selection group. An empty selection resets the objects selected (Fig. 7).

Navigation Interaction
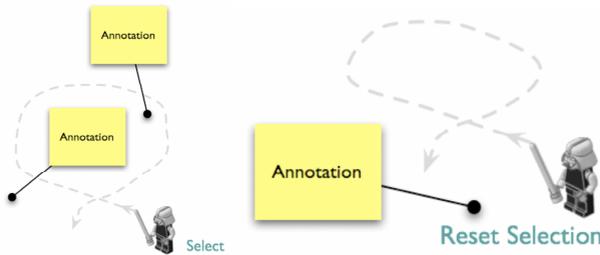
Fig. 7. Object selection and reset using laser interaction

Context Menu. Some options are only applied under a certain context. To satisfy this need, context menus automatically popup when a one or more objects become selected (Fig.8).

Rendering: The aim of the first prototype of the rendering component was to extend the OpenSG scenegraph library to support Precomputed Radiance Transfer. This first implementation was rather designed to get a correctly working implementation of the algorithm up and running than being tuned for high performance. This approach resulted in a CPU implementation of the run-time calculations of the PRT algorithm which was relatively easy to debug. However, doing the convolution of lighting and transfer function on the CPU prevents one from using display lists or vertex buffers to accelerate the rendering. The results of the lighting calculation have to be transferred to the graphics board every frame which eliminates the possibility of caching. For the second version of the renderer we have moved the convolution to the graphics board which significantly increases the performance of the renderer. Further more we have worked on moving the PRT preprocess from a per-vertex basis into texture space (which is not finished yet) and we have implemented a full HDR rendering pipeline which includes HDR reflection maps. The following sections give some detail about these improvements.
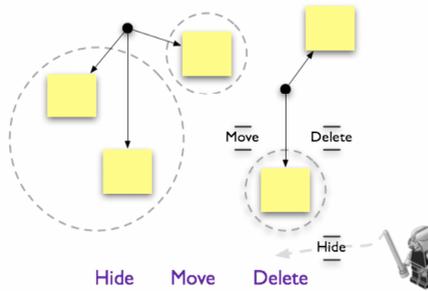

Fig. 8. Context Menu appears whenever a new object is selected. The available options depend on the selected objects type.

PRT Runtime calculation on GPU. The run-time calculation of the PRT algorithm boils down to a simple dot product between the coefficient vectors of SH-projected lighting and transfer functions. To perform this calculation using the GPU it is necessary to supply the per-vertex coefficient data to the graphics board. This is usually done using OpenGL vertex attributes, a general notion of data associated with a vertex. However OpenSG does not currently support the abstract vertex attributes, but only standard per vertex data like texture coordinates. Therefore we store the SH coefficients as texture coordinates. OpenSG supports up to 8 texture coordinate sets. We use up to 6 of these sets for the SH coefficients and also the secondary color property of a vertex, which allows us to support PRT rendering with up to five bands (i.e. 25 coefficients per vertex, stored in 6*4 texture coordinates + 1 secondary color coordinate). The first two texture coordinate sets are not used since they are used for standard texture mapping. The last information that is necessary to compute the PRT lighting on the GPU is the coefficient vector of the environment light. Since this vector is constant for the whole object we don't provide it as vertex attributes but as uniform variables. The coefficient vectors of the light and the transfer function at a vertex can now be scalar multiplied in a vertex shader. However, this implies that we

cannot use the simple OpenSG materials anymore but have to apply custom vertex and fragment shaders to the model. Fortunately the VRED editor does also support SHL materials (i.e. shaders written in the OpenGL shading language), so we can still use this tool for material application.

High Dynamic Range reflections. The OpenSG version we used during the development of the first prototype (version 1.6) did not support floating point texture formats and offscreen render targets. This prevented the use of high dynamic range reflection maps as well as the direct visualization of the light probe as the background of the scene. However in the current version of OpenSG (1.8) support for floating point textures and high dynamic range image formats (namely RGBE (.hdr) images) has been implemented. This allows us to use the light probe of a scene directly in shaders. A comparison between low dynamic range reflections and backgrounds and high dynamic range reflections and background is shown in Fig. 9. It should be noted that the only difference between the two renderings is the texture image format.
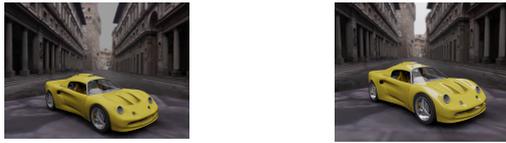
Fig. 9. Comparison of Low dynamic range reflections (left) and high dynamic range reflections (right)

5   Improve System Hardware

HMD: For the mobile outdoor and indoor design review sessions Trivisio has developed a first prototype HMD using OLED Technology. The first prototype is using an evaluation board from MED to drive the OLED QVGA microdisplays. To connect to VGA signals from a PC an external converter is used.

To allow for marker-less tracking we have integrated a VRmagic model "VRmC-3". This camera has a resolution of VGA (640x480) and offers progressive scan which is useful for tracking purposes. The refresh rate is up to 48fps. In the standard configuration the camera offers a field of view of 52° diagonal which seems to be suitable for a tracking camera and can be changed by replacing the lens. The interface is USB 2.0 which is standard for such cameras. Trivisio and IGD have experience with this type of camera from former projects. The mechanical design is only a first concept and the aesthetic aspects should be improved in the next steps. The focus is
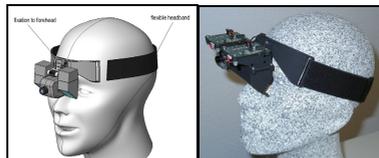

Fig. 10. HMD, fixation points to user's head

on functionality. The fixation to user's head is realized by a flexible headband as discussed and accepted with the projects's users and as this is the experience of Trivisio for the best and most ergonomic solution (see Fig. 10). In the case of calibration problems some additional sensors must be integrated or additional head fixation points such as nose-bridge must be added to the second prototype.

The two optical units are containing the optical components as described above including the microdisplays. The top covers can be removed and give an easy access to replace the microdisplays. Both units can be adjusted by a central screw to adapt the individual user's interpupil distance (IPD) by a linear sliding movement. The hinge connects the optical units to a forehead support and its angle can be adjusted to the individual user's forehead. The hinge can also be used to flip up the HMD to unblock user's sight quickly. Currently the FOV (field-of-view) is: virtual image: 17,6° diagonal, 10,66° vertical, 14,16° horizontal.

Powerwall: The Powerwall is a Tiled Display System (see Fig. 11). In this setup, the projection screen is divided in a number of tiles. Each tile receives the projection from on or two projectors (two in the case of stereoscopic walls). Each projector is individually connected to the video output of a single PC. All the PCs are arranged in a cluster configuration. The system additionally has a server, which role is to subdivide the image into individual image parts, and to send the image parts to the PCs over a network (usually using TCP/IP).
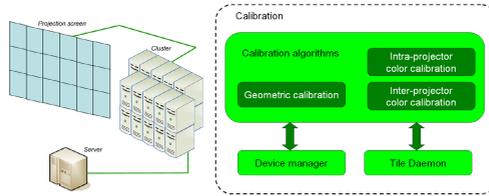

Fig. 11. The Powerwall

For the first prototype we used the HEyeWall from the Fraunhofer IGD in Darmstadt, Germany for the Powerwall component. The HEyeWall is a stereoscopic Powerwall, made out of 24 tiles (arranged in a 6x4 grid). Each tile is a part of the screen which receives the projection of two projectors (one for each eye). The total number of projectors is thus 48. Each projector connects to a PC, and the 48 PCs are used as a cluster of PCs. One supplementary PC serves as general server. Inside one tile, the right-eye and left-eye images are overlapping. When used in a monoscopic configuration, both projectors show the same image, resulting in an improved brightness. In the stereoscopic configuration, the images produces by each projector correspond to slightly different camera poses. The right-eye and left-eye signals are separated by specific goggles using Infitec Filters. The overall resolution of the HEyeWall is 6144 x 3072 pixels, for an image size of about 6m by 3m.

The calibration module is the software application used to calibrate the Powerwall. This application has three main components: The Calibration algorithms, the Device manager and the Tile Daemon (see Fig.11). The Calibration algorithms component is software that runs on the server of the Powerwall. It is used to conduct one of the

three possible calibration processes: the geometric calibration calculating the distortion compensation, the intra-projector color calibration using software shading table adaptation and the inter-projector color calibration doing the characterization of the projectors and then the color transformation so all tiles look the same.

6 Conclusion

We have presented the IMPROVE application scenarios and details of almost each individual component of IMPROVE. Additional components are marker-less and marker-based tracking as well as stereo video-streaming to allow for remote rendering of high quality images and wireless transfer to the HMD. The Communication Backbone linking all parts of the system has been presented as well as our interaction approaches with TabletPCs, Powerwalls and HMDs. High Dynamic Range Images are used for calculation of lighting, soft shadows and reflections. We have presented a new HMD and Color Calibration for the HEyeWall. We plan to continue our development of the IMPROVE system based on the OpenSG framework. In a separate paper also within this conference we discuss results of the first user tests.

References

1. [Regenbrecht02] Regenbrecht, H., Wagner, M.: Interaction in a Collaborative Augmented Reality Environment. In: Proc. Of CHI 2002, Minneapolis, Minnesota, USA, April 20-25, 2002, pp. 504–505. ACM Press, New York (2002)
2. [Schmalstieg06] Schmalstieg, D.: Designing Immersive Virtual Reality for Geometry Education, IEEE Virtual Reality, Alexandria, VA, USA, March 27, IEEE 2006. Virtual Concept 2006 Short Article Title Paper Number -11- Copyright Virtual Concept (2006)
3. [Klinker02] Klinker, G., Dutoit, A.H., Bauer, M., Bayer, J., Novak, V., Matzke, D.: Fata Morgana – A Presentation System for Product Design, In: Intl. Symposium on Augmented and Mixed Reality, Darmstadt, ISMAR 2002 (2002)
4. [BernEppstein03] Bern, M., Eppstein, D.: Optimized color gamuts for tiled displays. In: SCG '03. Proceedings of the nineteenth annual symposium on Computational geometry, San Diego, California, USA, pp. 274–281. ACM Press, New York (2003)
5. [Brown05] Brown, M., Majumder, A., Yang, R.: Camera-Based Calibration Techniques for Seamless Multi-Projector Displays. IEEE Transactions on Visualization and Computer Graphics, vol. 11(2) (2005)
6. [Debevec97] Debevec, P.E., Malik, J.: Recovering High Dynamic Range Radiance Maps from Photographs, Computer Graphics, vol. 31, pp. 369–378, In: SIGGRAPH 1997 (1997)
7. [Debevec98] Debevec, P.: Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In: SIGGRAPH 98 (July 1998)
8. [Drummond02] Drummond, T., Cipolla, R.: Real-Time visual tracking of complex structures. IEEE Transaction on Pattern Analysis and Machine Intelligence 27, 932–946 (2002)

9. [Fiorentino02] Fiorentino, M., de Amicis, R., Monno, G., Stork, A.: SpaceDesign: A Mixed Reality Workspace for Aesthetic Industrial Design, In: Intl. Symposium on Augmented and Mixed Reality, Darmstadt, ISMAR 2002 (2002)
10. [Fröhlich04] Fröhlich, B., Hoffmann, J., Klüger, K., Hochstrate, J.: Implementing Multi-Viewer Time-Sequential Stereo Displays Based on Shuttered LCD Projectors, 4th Immersive Projection Technology Workshop, Ames, Iowa, May 2004 (2004)
11. [Kato00] Kato, H., Billinghurst, M., Poupyrev, I., Ikamoto, K., Tachibana, K.: Virtual object manipulation on tabletop AR environment. In: International Symposium on Augmented Reality, 2000, pp. 111–119 (2000)
12. [Kresse03] Kresse, W., Reiners, D., Knoepfle, C.: Color consistency for digital multi-projector stereo display systems: the HEyeWall and the Digital CAVE. In: EGVE '03. Proceedings of the workshop on Virtual environments 2003, Zurich, Switzerland, pp. 271–279. ACM Press, New York (2003)
13. [Lepetit05] Lepetit, V., Fua, P.: Monocular Model-Based 3D Tracking of Rigid Objects: A survey. (October 2005)
14. [Lowe01] Lowe, D.: Distinctive image features from scaleinvariants keypoints. International Journal of Computer Vision 20(2), 681–682 (2001)
15. [Majumder00] Majumder, A., He, Z., Towles, H., Welch, G.: Achieving color uniformity across multi-projector displays. In: VIS '00. Proceedings of the conference on Visualization '00, Salt Lake City, Utah, United States, pp. 117–124. IEEE Computer Society Press, Washington, DC, USA (2000)
16. [Majumder03] Majumder, A., Jones, D., McCrory, M., Papka, M.E., Stevens, R.: Using a Camera to Capture and Correct Spatial Photometric Variation in Multi-Projector Displays, In: IEEE Workshop on Projector-Camera Systems (PROCAMS 2003) (2003)