Int J Interact Des Manuf (2007) 1:115–126
DOI 10.1007/s12008-007-0010-8

TECHNICAL PAPER

IMPROVE: An innovative application for collaborative mobile mixed reality design review

Pedro Santos · André Stork · Thomas Gierlinger · Alain Pagani · Céline Paloc · Iñigo Barandarian · Giuseppe Conti · Raffaele de Amicis · Martin Witzel · Oliver Machui · José M. Jiménez · Bruno Araujo · Joaquim Jorge · Georg Bodammer

Received: 13 September 2006 / Revised: 10 October 2006 / Accepted: 17 October 2006 / Published online: 16 May 2007

Abstract In this paper we introduce an innovative application aiming at combining large, tablet-based and head-mounted displays for 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 Elasis FIAT. Our activities are supported by the EU IST project IMPROVE. It covers activities in the areas of: HMD development using unique OLED technology, markerless tracking, augmented 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 test.

Keywords Mixed reality applications · Collaborative Design Review

1 Introduction

Design Review is one of the most common applications of Virtual Reality and Immersive Projection Technology within industry. A typical use scenario looks like follows: many observers are sitting in front of a large projection wall viewing and discussing a virtual model. New developments allow up to four tracked users [1].

In the last couple of years Augmented Reality (AR) also found first applications in design and design review (e.g., see ARVIKA project homepage: www.arvika.de). Typically

P. Santos (✉) · A. Stork
Fraunhofer-IGD, A2, Darmstadt, Germany
e-mail: Pedro.Santos@igd.fhg.de

A. Stork
e-mail: Andre.Stork@igd.fhg.de

T. Gierlinger · A. Pagani
Tu-Darmstadt, Darmstadt, Germany
e-mail: Thomas.Gierlinger@igd.fhg.de

A. Pagani
e-mail: Alain.Pagani@igd.fhg.de

C. Paloc · I. Barandarian
Vicomtech, San Sebastian, Spain
e-mail: Celine.Paloc@vicomtech.es

I. Barandarian
e-mail: Inigo.Barandarian@vicomtech.es

G. Conti · R. de Amicis · M. Witzel
Graphitech Trento, Italy
e-mail: giuseppe.conti@graphitech.it

R. de Amicis
e-mail: raffaele.de.amicis@graphitech.it

M. Witzel
e-mail: martin.witzel@graphitech.it

O. Machui
Trivisio Dreieich, Germany
e-mail: machui@trivisio.de

J. M. Jiménez
STT San Sebastian, Spain
e-mail: jmjimenez@stt.es

B. Araujo · J. Jorge
INESC-ID, Lisbon, Portugal
e-mail: brar@immi.inesc.pt

J. Jorge
e-mail: jaj@inesc-id.pt

G. Bodammer
Microemissive Edinburgh, UK
e-mail: georg.bodammer@microemissive.com

video-based AR is applied to augment physical models with design alternatives e.g., new virtual dash boards (VW), alternative wheel designs (Audi). Interaction and collaboration seldom take place in such applications.

SpaceDesign [2] 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 [25] 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 [3] 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 in presentation facilities.

For us this was the starting point for brainstorming the possibilities of AR 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 extend 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.

Fascinated by the potential of mobile collaborative AR, the architects strive for the following scenarios (Fig. 1):

1. The design of buildings begins with early shape studies aiming at integrating the new building in its surroundings and the landscape—the architects wish to see this stage supported 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. Then, they want to take their sketch home into their office for refinement. The result shall be reviewed collaboratively and presented to customers in an indoor environment using a couple of HMDs to support multiple users with individual viewpoints. Pen-based direct 3D interaction is envisaged for creating annotations and change orders.
3. Finally, when everything is done, the architects want to bring the final design back to the physical construction site for a multiple user presentation. Again, many 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.

In contrast to the architects, whose scenarios are centred around the use of HMDs, the representatives from automotive industry are familiar with large area displays and want to improve and extend the use of large projection technology. The automotive scenario looks as follows (Fig. 2):

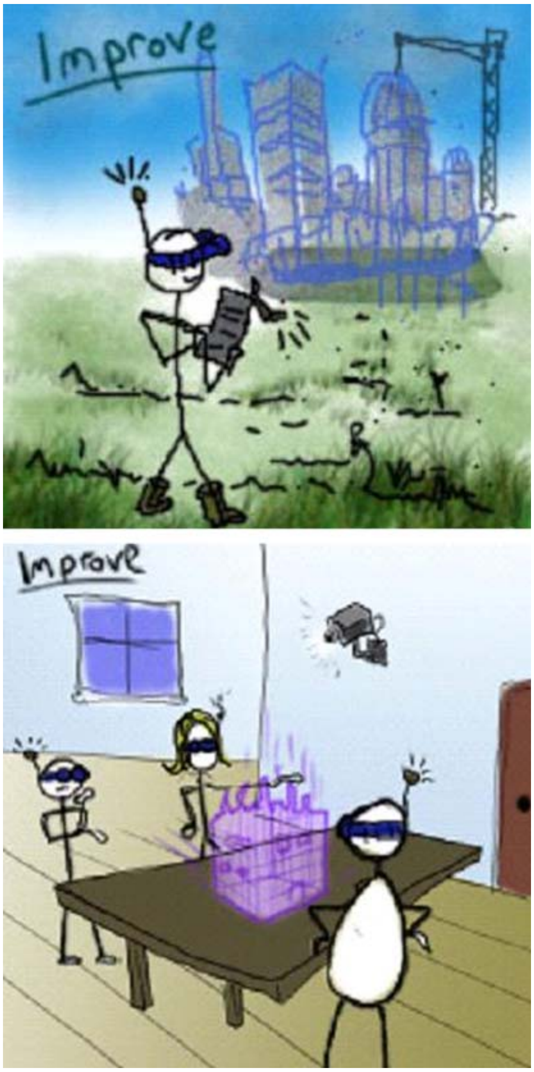

Fig. 1 Onsite sketching and collaborative indoor reviewing scenarios (conceptual sketches)

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 color consistency of high resolution walls more efficient calibration techniques are required.

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 seethrough 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 [26] while IMPROVE is wider in scope and spans more diverse environments and scenarios. IMPROVE not only contributes to the mixed and AR 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.

The paper first introduces the system design chosen to realize the scenarios, then it presents the developments in the various fields. The paper concludes with results from first user test and a look into future developments.

2 System design

The IMPROVE system is build upon five distributed, autonomous subsystems: Communication Backbone, Interaction, Tracking and Rendering Component and a Central Repository. Each subsystem communicates with the other modules through a high-level xml-based message exchange. This choice provides IMPROVE with a very high level of flexibility in terms of physical distribution, ease of deployment and robustness. In fact since protocols are consistently defined each component becomes both independent and replaceable. Further the number of IMPROVE clients can be scaled up arbitrarily if kept within the computational capacity of the chosen communication backbone server. The Communication Backbone connects through XML messages to provide the required data exchange mechanisms. Messages can be categorized into modelling, annotation, configuration, calibration and shape data. Each software component registers to which data they are interested in by subscribing to various channel topics each defining a specific data set. The interaction component takes care of model/scene/session management, shape management, view management, annotations, design modifications, light creation and control and it follows a publisher/subscriber approach. Interactions can be distinguished into two different kinds: the first is applied locally; the second is additionally transmitted via the communication backbone to the other components. The tracking component sends the information captured by the different devices supported and it wraps them into normalized XML messages called “IMPROVE Entities” (IE) and it sends this data via the client channel through the communication backbone. The rendering component visualizes the virtual models and annotations using a variety of display devices. Finally the repository implements the functionality to store and retrieve invoked commands and it stores configuration data like calibration, meta-information and setup data (Fig. 3).

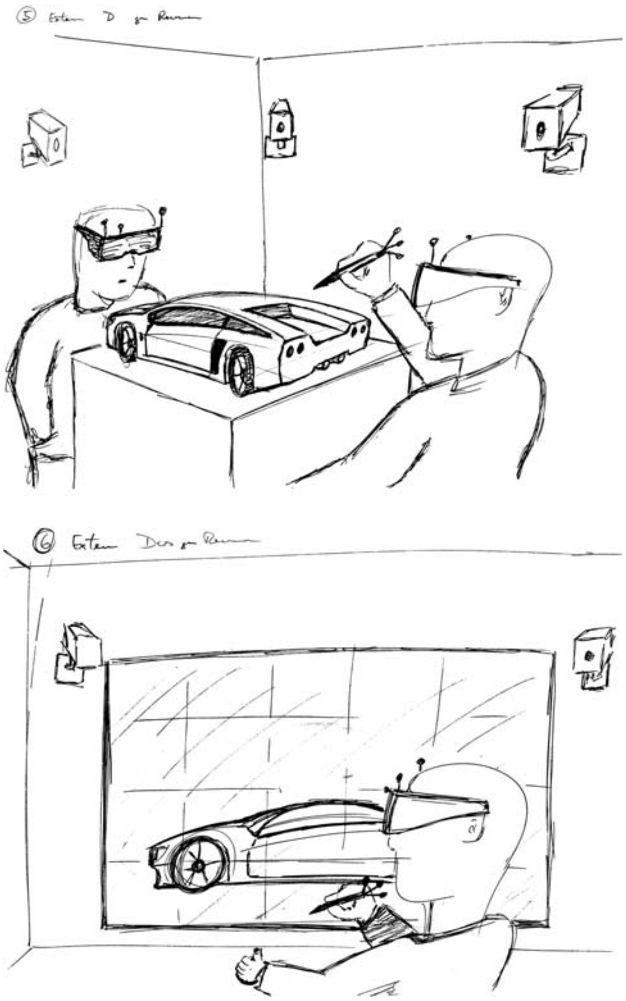

Fig. 2 Collaborative cAR design review scenario combining large and head mounted displays (conceptual sketches)

3 The OLED-based IMPROVE HMD

The IMPROVE consortium comprises the companies MED (the world record holder for the smallest display) and TriVisio (well known in the AR community for their HMDs). Together

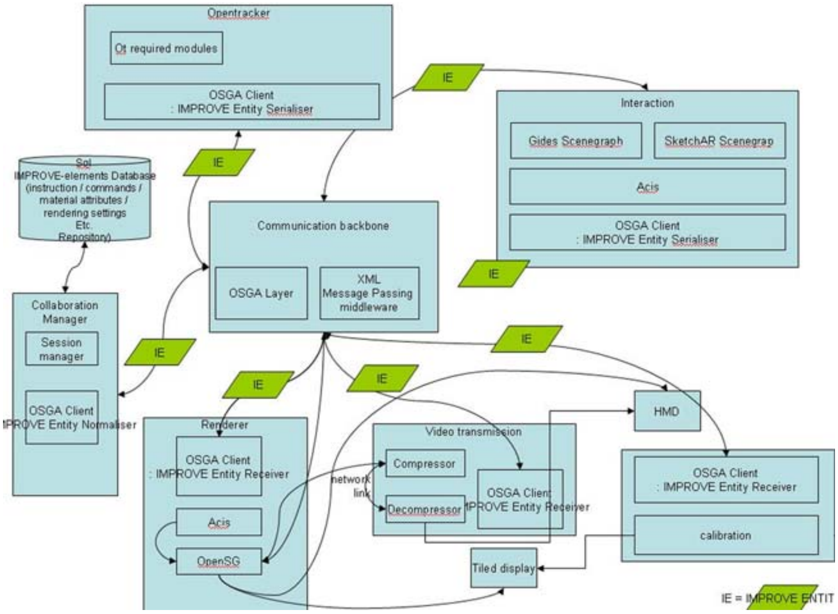

Fig. 3 System design overview

they develop a new stereoscopic optical see-through HMD focussing especially on the comfort and acceptance of the HMD. MED can control the production process for their OLEDs and improve the brightness level of their OLEDs for outdoor use. OLEDs are well suited for HMDs due to the following characteristics: they offer super-fast blur free technology, an ultra-compact footprint, low power consumption (no external batteries), light output proportional to current driven through the display and Lambertian light emission with almost no viewing angle restrictions.

To meet the outdoor brightness requirements, there are two ways to increase luminance of an OLED display:

1. Generate more light inside the device,
2. Extract more of light that is generated (OLEDs are typically only 20% efficient at light extraction).

MED has focused on increasing luminance using both of the above strategies with significant degree of success. Optimising the interface properties internally to the OLEDs shows encouraging results in this area. To extract more light we can optimise layer thicknesses to enhance constructive interference from the optical cavity structures that exist in the device. Layer thickness on light extraction in an OLED microdisplay device has been modelled and experiments to complement the modelling work have been carried out. These experiments show that predicted luminance can be increased from 900 to approx. 2,400 cd/m^2.

With the novel MED proprietary developments a three fold increase in luminance output based on the current emitting materials can be achieved.

The MED OLEDs build the basis for a new stereoscopic optical see-through HMD designed by TriVisio (Fig. 4). The design is aiming at the following criteria:

- low weight,
- high transparency (unoccluded view to the real world—no tunnel view),
- large opening angle—no need to adjust inter-pupil distance,
- no need for additional external power supply–complete power supply through USB,
- passive stereo support though single video cable.

Chosen HMD concept offers the maximum see-through brightness and minimum blocking of user's sight. The weight is less than 200 g, the field of view 17.6° diagonal and the exit pupil diameter is 8 mm. Future development steps foresee the extension towards a pixelwise real-light blocker and the use of the brightness extended OLEDs.

4 Tiled display colour calibration

Large multi-projectors displays have gained considerable importance in the last decades as scalable, inexpensive and high-precision visualization systems. Their extremely high resolution makes them particularly attractive for collaborative work and industrial design. However, achieving the illusion of one large seamless display relies on a precise calibration of the system. This calibration aims to solve three specific problems [4]:

- The variation of the geometric properties of the individual display units,
- The intra projector color variation (color variation within a single projector) and
- The inter projector color variation (color variation across different projectors).

Fig. 4 Trivisio optical see-through HMD

Even if satisfying solutions have been proposed for the geometric calibration problem ([5–7] to cite a few), the color calibration of a tiled display remains an open issue, which has been only partially addressed by using different approaches [7]. The Gamut matching approach [8–11] assumes that the intra-projector variations can be neglected, and the color gamut of each projector (i.e., the set of colors the projector can display) is measured. A common gamut is then computed and a conversion from a specific gamut to the common gamut is done using linear transformations. The main limitations of this approach is the case of strong intra-projector variations, and the high time consumption. The Intensity manipulation approach [4,6,12–14] assumes that the chromatic properties of the projectors are uniform across the display. In this special case, the color variations are only due to varying intensities of the different color channels. The display's intensity response function is first modelled and measured, and the input image is corrected using the inverse model to compensate the photometric variations. However, this approach excludes using different models of projectors in the same setup (Fig. 5).

Within the scope of IMPROVE, we developed a new method to calibrate a tiled display. This method mixes the advantages of the gamut matching and intensity manipulating approaches in a two-step process. With the tenable assumption that the chromatic properties of each channel of a single projector are spatially invariant, we developed an iterative algorithm for intensity compensation of one projector, thus solving the intra-projector variation problem. We then solved the inter-projector variation problem by deriving a fast algorithm for finding a common gamut between n projectors in time O(n). The main benefit of our method is the generality of the calibration technique. Our fast gamut matching algorithm allows for tiled displays with large chrominance shifts (with e.g., projectors from different vendors), and our iterative shading correction does not make the assumption of a spatially invariant intensity transfer function for each projector.

Fig. 5 HEyeWall: High resolution stereoscopic tiled display with 48 beamers (courtesy of Fraunhofer IGD)

4.1 Iterative shading correction

The main idea of the method is to modify the intensity of the input pixels to compensate the original intensity differences. Almost all known methods for intensity manipulation define a parameterized model of the intensity variation, measure sample points of this model and apply the inverse of the model for correction. The disadvantage of model-based solutions is that they rely on simplifying assumptions to reduce the complexity of the model. For example, they often stipulate that the normalized intensity transfer function of a projector does not vary spatially.

To reduce the number of assumptions, we opted for an iterative loopback call method with progressive input manipulation. We now want to find the best correction value for each level of each point to achieve color uniformity for a projector. To this aim, we use a common digital camera as a luminance measuring device. Because we use a camera with adjustable exposure time, we can augment the dynamic range of the camera by taking the measurements in high dynamic range (HDR) images [15]. A preliminary geometric registration is performed using a homographic transformation of the image.

To work around the problem of global brightness instability, we developed a method based on instant comparison between displayed points at different spatial locations. To this aim, we define a set of points with similar measured luminance as target points, and measured the luminance of the other points relatively to the target points. The iterative algorithm then modifies the input value for a given channel

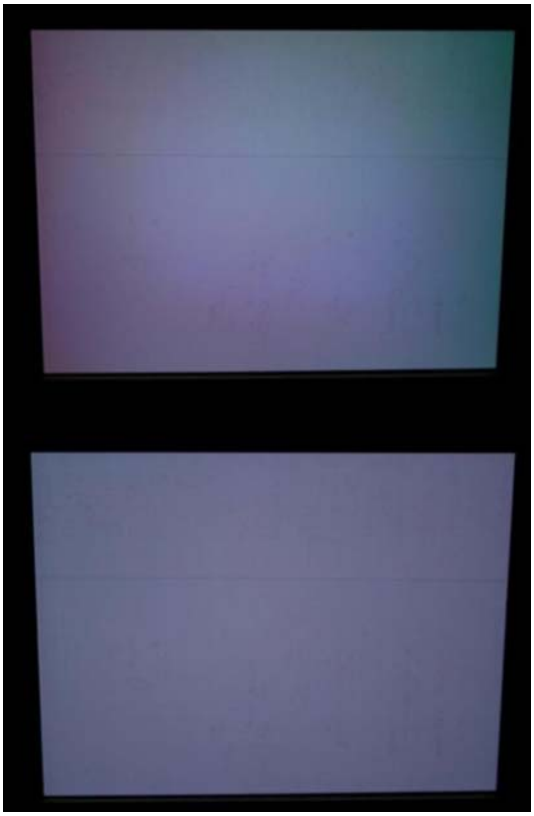

Fig. 6 A projector before (*top*) and after (*bottom*) intra-projector correction

to brighten dark regions and darken bright regions towards the intensity of the target points.

4.2 Fast gamut matching

Known explicit methods for gamut matching [10,16] have the disadvantage of having a high complexity.

(O(n^3)), preventing their use for large system with 40–50 projectors. Instead of seeing gamut matching as a geometrical optimization problem, we find an approximate optimal gamut by growing an initial set of colors inside the set of commonly displayable colors. To this aim, we developed a method to efficiently find out if a given color is displayable by all the projectors, and seek for the biggest region of colors having this property. Our algorithm gives near perfect results with a significantly reduced complexity. Indeed, we can compute an approximate common gamut between n projectors in time O(n). Once a common gamut has been found, the input values for each projector can be linearly corrected to fit in the common gamut (Fig. 6).

5 Mixed reality rendering techniques

When performing realistic rendering in mixed reality environments our goal is to seamlessly integrate virtual objects into a real scene. This has to be done in real-time, so that the user can move freely and examine the virtual objects from arbitrary viewpoints. Besides the general problems of realistic rendering like accurate material simulation and efficient calculation of global effects (reflections, shadows, caustics) we also face the problem of consistent lighting. This means that the virtual objects should respond to the same lighting as the real ones—shadows should be cast into the same direction as those of real objects and the colour and intensity of the light sources that light the virtual model should approximate the real lighting environment as close as possible. To solve these problems we use a combination of the following three techniques:

1. Image Based Lighting
2. High Dynamic Range Imaging
3. Pre-computed Radiance Transfer

Image based lighting (IBL) is a technique introduced by Paul Debevec [17] to enable consistent lighting of virtual objects. The idea is to take a photograph of a lighting environment (usually an environment map by photographing a mirror sphere or by using a special panoramic camera) and use the pixel values in this image as incoming radiance from that direction. This technique assumes that the lighting environment is far away from the object that is lit, which is acceptable for both usage scenarios: In the architectural scenario we are usually interested in lighting a virtual building with a light probe that represents the sky. In the automotive scenario we want to review a car in a showroom, where the lights of the showroom (windows or light source at the ceiling) are also relatively far away from the car itself.

High dynamic range imaging (HDRI) [15] is used to increase the accuracy of the IBL approach. Usual cameras are only able to capture a limited dynamic range of the incoming light. Depending on the exposure settings of the camera overexposed parts of a photograph are clamped to white (loosing all structural information in these parts) while underexposed parts are plain black. HDRI solves this problem by using a series of photographs shot at different exposure levels instead of a single one. The information from the photographs of the series is then combined so that meaningful lighting information is available for the dark parts of a scene as well as for the very bright ones.

Pre-computed radiance transfer (PRT) was introduced by Peter-Pike Sloan [18]. It is a technique that allows real-time rendering of global effects like shadows, caustics and colour bleeding in low frequency lighting environments. The basic idea is to transform the problem of lighting calculation into

Fig. 7 Standard OpenGL lighting

Fig. 8 Pre-computed radiance transfer—note the occlusion effect in the interior and the smooth shadows

a different domain where the necessary operations can be performed more efficiently. The main problem when solving the rendering equation is to calculate the integral over incoming light directions. PRT solves this problem by projecting the involved functions (i.e., incident lighting and cosine-weighted visibility, which is called the transfer function) into the Spherical Harmonics basis. For diffuse materials the integration over the hemisphere reduces to a dot-product which can be performed in real-time. The projection of the transfer function is done in a pre-process, while the convolution with a lighting environment is done at runtime. This technique has the advantage over other precomputation techniques like radiosity in that the lighting environment can be dynamically changed without the need to redo the pre-process.

The rendering component is based on the OpenSG (www.opensg.org) scenegraph. This scene graph is open source and has support for different display devices like tiled displays and head mounted displays already built in. The projection of the transfer function is done in a pre-process on the vertex level and the resulting coefficients are stored as so called attachments together with the geometry. The runtime calculation is currently done on the CPU, where a new vertex colour is computed depending on the current lighting environment and orientation of the object. Figures 7 and 8 show the difference between standard OpenGL lighting and PRT (Light probe courtesy of www.debevec.org, car model courtesy of dmi.chez-alice.fr).

6 Indoor/outdoor large area tracking

The main goal of the Augmented Reality technology is to add computer-generated information (2D/3D) to a real environment in such a manner that the real and virtual objects appear as co-existing in the same world. In order to get a good illusion, the registration problem must be addressed. In this way, the position-orientation (pose) of the camera respect to a reference frame must be accurately estimated or updated over time. In this work, we address the registration problem for interactive outdoor AR applications, working on a fully mobile wearable AR system that combines a vision-based (optical) tracker with an inertial tracker.

Though the real-time registration problem using computer vision techniques has received a lot of attention during last years, it is still far from being solved [19]. Ideally, an AR application should work without the need of adaptation neither the object nor the environment to be tracked, by placing special landmarks or fiducials [20]. This issue is known as markerless tracking.

We can divide the optical markerless tracking technology in two main groups: model-based tracking and move-matching. In model-based tracking, the identification in the images of some features (texture patches or edges) corresponding to a known 3D model is used to solve the registration problem and therefore estimating the camera pose [21]. Though this type of tracking can achieve good results, it can be unsuitable for many applications such as outdoor tracking, because of the need of a model that should be constructed beforehand, which can be a very time consuming task. The move-matching approach, also known as structure from motion, tends to estimate the pose of the camera related to a 3D structure also obtained from the imaged scene, without a priori knowledge about the scene, by tracking (matching) features [27,28] extracted from the images over time. These methods can achieve very accurately results but using some very time consuming techniques like bundle adjustment, that are not suitable for interactive applications.

We are working on an optical markerless tracking method which combines the advantage of move-matching and model-based tracking. Our approach is based on planar surface tracking [22]. In an outdoor scenario, the ground and the building façades can be seen as planes. These 3D planes allow us to recover the camera pose by tracking natural local features extracted from that surface. When starting the tracking process, the initial planar projective transformation (homography) between a world plane and its projection in the image must be estimated by selecting manually four coplanar points. The camera pose can be extracted from the homography by knowing the internal camera parameters that should be calibrated beforehand [23]. The next camera poses can be calculated by chaining the homographies estimated from the consecutive views of the same planes. These homographies

Fig. 9 Homography tracker test-run at Plaza de San Telmo, San Sebastian, Spain

can be accurately estimated by matching feature points corresponding to inlier data. These inlier data can be obtained by using robust methods [24]. In this way, we are studying different robust estimation algorithms used in computer vision, such as RANSAC or M-estimators, and their complementariness. These estimators allow the tracking process to be robust against noise or spurious data (outliers) produced by mismatching (Fig. 9).

It is well known that the recursive tracking methods, like chaining homographies, can present some drift after a period of time because of error accumulation, even with robust estimation [19]. In this way, we are also using tracking information given by an inertial tracker in order to correct the deviation of the optical tracking. Fusing the two tracking subsystems gives us the benefits of both technologies, while the sensor's complementary nature helps overcome sensor-specific deficiencies.

7 Collaborative interaction for sketching and design review

The visit to the industrial partners revealed serious limitations of the tiled display systems. In particular, they only use these kinds of display only as large screens and usually the content which viewed is controlled by an operator which is seated back in the room controlling the viewing. There is no ability from the participants of a review session to interact directly with the content. Usually they communicate with the operator through speech to control objects or the view. Of course, this approach limits the usage of the Tiled Display system (TDS) in design review process and the benefits of these devices in their workflow. However, even if most of the time it is used as showroom for both architecture and automotive industry, it is still a complement and the scale brings realism to the content. The same can be said regarding HMD; usually they are used as a simple alternative visualization tool. In order to take advantages of user sketching skill, we choose to base our interaction in the usage of TabletPC computers to interact with both TDS and HMD devices (Fig. 10).

Fig. 10 Architectural scenario using HMD in VR/AR environment

7.1 Interaction with TabletPC for architectural scenarios

Figure 10 presents an overview of the architectural scenarios and setup. On these scenarios, the user interacts with the system using a TabletPC. The user is able to view the 3D content and its workspace using Head Mounted Display. In order to implement the customized functionality for the architectural scenario, GIDES++ is integrated within IMPROVE and further developed. GIDES++ is a 2D sketching interface able to offer basic CAD functionality taking advantages of sketch based interaction. We decide to re-use this interface as platform for the innovative interaction, since it presents a simple modelling tool using Spatial ACIS kernel and support sketching interaction which is adapted to the architectural user taking advantages of their sketching skills. On the other hand, this system is able to run on TabletPCs allowing sketched annotations and offers the ability to present natural interface mimicking the pencil/paper metaphor.

7.2 GIDES++ system for collaborative design review

The GIDES++ system is used as an interface to access to IMPROVE functionality and will interact with the other IMPROVE sub-systems such as Rendering, Video Transmission and Tracking through the usage of the IMPROVE communication backbone. Thanks to the interaction with the other sub-systems, we are able to handle collaborative scenarios where several users are able to interact or making design review using HMD or TDS. Figure 11 depicts those scenarios. On both scenarios, architects are able to interact with the IMPROVE system though the usabe of TabletPC computers running our system. To support these scenarios, we implement in GIDES the ability to work in colaborative envorinments sharing modelling and object manipulation

Fig. 11 Collaborative design review scenarios

ability through the OSGA package. More details about this implementation will be presented below.

7.3 Multimodal modalities and view/object controls

The system enable the user to control the view and objects orientation through a simple widget following the three axis representation of orthonormal basis. This widget allows to control the view, however this representation is limited for 3D view control. We also implemented an independent view/object control based on the usage of a space mouse device. The space mouse was integrated to GIDES using the OSGA package, doing so, all the control events of the device (i.e., transformation of the space mouse) were sent to GIDES as messages and the view and object were updated. As alternative, some voice commands are implemented in GIDES to control view and some of the modelling funcionality.

7.4 Annotation system for architectural scenario

We implemented in GIDES++ framework two different kinds of annotations: for objects and for views. The object annotation type enables users to comment and discuss a specific area of an object, giving them a real sense of 'post-it' in form of simple sketch. By clicking on an object location a popup/panel appears allowing users to draw/sketch annotations concerning this point. There is also a simple mechanism for displaying and hiding these annotations by simply clicking a widget that appears at the point where the user clicked. This point is defined by a marker that permits to identify and retrieve an annotation. One GIDES++ object could have several markers associated to him. The annotations window appear as billboard, meaning it will always face the camera due GIDeS++ has an orthogonal viewpoint (Fig. 12).

The annotation data is composed of the position of the marker, position of the window annotation, and a dataset of strokes that describe the sketch drawed. All of this information is saved with the model and will persistent along the application. The annotations for views will present users a translucent whiteboard on top of the scene. They will allow for greater freedom for review a design, and permit user to annotate information in large scale, that permits users transmit ideas and comments while discussing the overall project. This feature could be very useful in collaborative interaction to review design process. This kind of annotation is supported on a transparent panel (enabled bitmap with an alpha channel) that covers the whole model on the screen. User can sketch over this panel and when he is finished a screenshot will be taken and added as a thumbnail to a menu at the bottom of the screen. We can create several annotations or views. When hovering over the several thumbnails info will appear in a popup and, by simply clicking on the image the user will be taken to the annotation's viewpoint and the sketch strokes will be newly displayed. For this kind of annotations only camera information is important as its size and position is set beforehand.

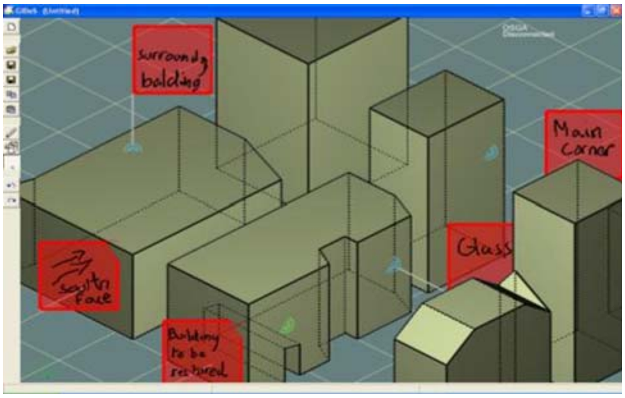

Fig. 12 Examples of annotations

7.5 Collaborative functionality supported

We have designed and implemented a collaborative mechanism to share the creation of simple shapes (such as spheres, cubes, cylinders, cones and pyramids) and geometric transformations of these objects. Based on communication backbone, we implemented a client that permits to publish and subscribe messages of the topics related to creation or update of shapes or geometrics transformations over them. So, when a user create a shape on this interface (application instance), the data information that describe this object is published to backbone communication system, and other clients that subscribed this kind of messages will receive that message, and a predefined function callback will be activated in order to create that shape in local system. In this way all the IMPROVE remote agents could synchronize the same shapes modifications in real-time. This communication architecture allows sharing shapes between several instances and allows visualizing its content on HMD or TDS devices. On the other hand, the annotation presented above are also integrated in a collaborative system that permits to share an annotation with other users, TDS based viewer or HMD devices. The

Fig. 13 Collaborative scenario with two users using TabletPCs in front of a Powerwall

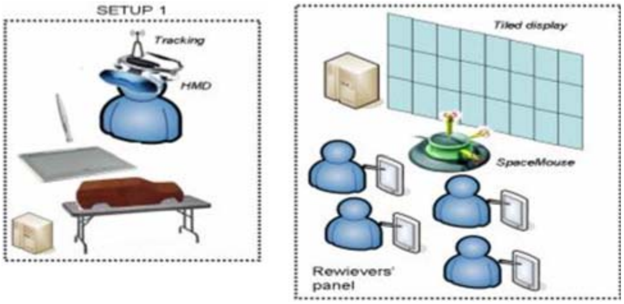

Fig. 14 System configuration for collaborative automotive design review

general collaborative mechanism of annotations is similar to the shapes or transformations. The following figure shows the schematic setup of the collaborative automotive scenario with many users in front of a display wall interacting with Tablet PCs and one (or more) users with HMD(s) beside (Figs. 13, 14).

8 Stereoscopic video transmission for mobile Users

In earlier sections of the paper we introduced the requirement of the users to move freely (un-tethered) through a larger area and stressed the need for augmented reality rendering techniques to embed virtual objects seamlessly into real scenes maintaining consistent shading and shadowing.

Algorithms which support this to a considerable extend, such as PRT, are computationally demanding. It is a matter of fact that the latest and fastest CPU and GPU technology is always introduced in desktop machines before being a available in mobile computers.

Therefore, we decided to leave the rendering up to the desktop machine(s) and apply and optimize video streaming technology for transmitting the rendered stereoscopic images to the mobile users (Fig. 15).

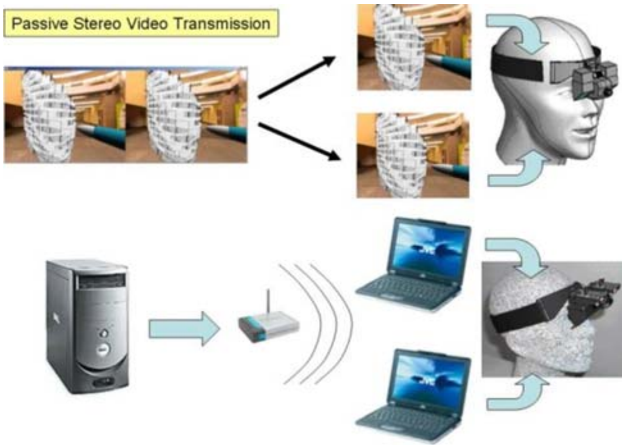

Fig. 15 Stereoscopic synchronized video transmission

As receiving machines we have chosen JVC subnotebooks that offer good CPU performance at a low weight—actually the display and the keyboard could be dropped to reduce weight.

We evaluated many state of the art components as basis for our video transmission module, such as:

- GDI, DirectX, MS Media SDK
- DivX, XviD, FFmpeg, HDX4
- Connection-oriented RTP layer over connection-less UDP, alternatively: proprietary connected-oriented TCP
- GDI

Our SW frame grabber component grabs an image with 1,024 × 768 pixels within 3 ms under PCIe. As codec we have chosen FFmpeg which is similar to XviD in performance, but allows to control more parameters and offers more optimization capabilities. As transmission protocol we have developed a proprietary connection-oriented TCP with cross synchronizing the two streams sent to the two receiving subnotebooks for image-synchronous display.

The cross stream synchronization synchronizes threads on sender side. After compression both current frames are synchronously transmitted over TCP. The synchronous streams are visualized on the receiver side. In case that frames need to be dropped to maintain real-time performance with minimal lag, cross-synchronized frames are dropped simultaneously on both channels.

The following Table 1 shows the performance of our solution measured on an Intel Pentium 4 3.0 GHz HT with Nvidia Geforce 6800 GTX PCIe Our solution is completely independent from any screen content or graphics subsystem used, it works with OpenGL, Direct3D and any other screen (frame

Table 1 Performance of video transmission component

Resolution ($w \times h$)	Framerate (fps)
800×600	64
1,024×768	45
1,280×1,024	30
1,600×1,200	21

buffer) content and is up to 15 times faster than commercial application sharing tools.

9 User evaluation

In this section we want to summarize first feedbacks of preliminary informal user tests

HMD:

+ “surprisingly” light
+ very bright—it almost blocked what was behind (without being solid)
+ very sharp image
+ comfortable to wear (for a prototype, for approx. 90 min)
+ menus readable—easy to manipulate
+ correct 3D impression—menus visible in the right depth
+ pictures immediately visible
− small field of view (due to size of MED OLEDs)
− too greenish (OLED prototypes)
+/− resolution ok—to be further evaluated

TabletPC interaction:

+ natural way of interaction
+ easy-to-use interface
+ immediate collaborative feedback—very good response time
+ handwritten annotations
− limited modeling functionality compared to commercial CAD tool
− menu too big (drag ‘n’drop 3D menus and overlays)

Immersive environment with passive stereo projection:

+ almost no tracking lag
+ very good 3D impression + sharp and good image quality
− 3D pen more difficult to control than TabletPC with 2D pen

Immersive environment with new HMD and video transmission:

+ almost no lag even with video transmission—annotations appear immediately
− contradictory depth information from proprioceptive versus visual system (HMD focal distance)

Rendering component (shadows and visual quality):

+ how shadows and materials react to light is impressive (PRT and HDR-based rendering)
+ much faster navigation than with commercial tools (e.g., Artlantis)
− PRT currently limited to static scenes

One of the things that need further evaluation for beneficial use in an industrial environment is the setup and preparation time, because ‘time is money’.

10 Conclusion and future work

In this paper we have presented a unique effort in hardware and software research and development to facilitate collaborative mixed reality design review in indoor and outdoor scenes with mobile users.

We have given an overview of the hardware-oriented developments in the field of organic light emitting and head mounted displays. We have introduced the software and system architecture of the IMPROVE prototype system and the capabilities of its various components. Plus we have summarized the feedback of first informal test with real users from the architectural and automotive field.

Future work is dedicated into the following directions:

- Brighter OLED and light blocker mechanism for the HMD
- Fully automatic calibration method for tiled displays (not only for intraprojector color correction but also for interprojector color correction)
- Extending mixed reality rendering techniques towards dynamic scenes and accelerating (pre)computation
- Faster and more robust markerless tracking
- Diversity of interaction techniques and extension of functionality
- GPU-based video en-/decoding

And last but not least we will do more user tests with users from Page/Park architects, Elasis Fiat and others.

Acknowledgements The authors want to acknowledge the EU IST programme for funding the IMPROVE project.

References

1. Klüger, K., Fröhlich, B., Hoffmann, J., Hochstrate, J.: Implementing multi-viewer time-sequential stereo displays based on shuttered lcd projectors. In: 4th Immersive Projection Technology Workshop, Ames, Iowa (2004)
2. Monno, G., Fiorentino, M., de Amicis, R., Stork, A.: Spacedesign: A mixed reality workspace for aesthetic industrial design. In: Symposium on Augmented and Mixed Reality, Darmstadt,ISMAR (2002)

3. Klinker, G., Dutoit, A.H., Bauer, M., Bayer, J., Novak, V., Matzke, D.: Fata Morgana—a presentation system for product design. In: International Symposium on Augmented and Mixed Reality, Darmstadt, ISMAR (2002)
4. Majumder, A., Stevens, R.: Color nonuniformity in projection-based displays: analysis and solutions. IEEE Trans. Vis. Comput. Graph. **10**(2), 177–188 (2004)
5. Beardsley, P., Willwacher, T., Rao, S., Raskar, R., van Baar, J., Forlines, C.: Ilamps: geometrically aware and selfconfiguring projectors. In: Proceedings of ACM SIGGRAPH (2003)
6. Majumder, A., Towles, H., Raji, A., Gill, G., Fuchs, H.: Pixelflex2: a comprehensive, automatic, casually-aligned multi-projector display. In: IEEE Workshop on Projector-Camera Systems (PROCAMS) (2003)
7. Majumder, A., Brown, M., Yang, R.: Camera-Based Calibration Techniques for Seamless Multi-Projector Displays, vol. 11 (2005)
8. Stone, M.C.: Color balancing experimental projection displays. In: Color Imaging Conference, pp. 342–347 (2001)
9. Stone M.C.: Color and brightness appearance issues in tiled displays. IEEE Comput. Graph. Appl. **21**(5), 58–66 (2001)
10. Chen, H., Wallace, G., Li, K.: Color gamut matching for tiled display walls. In: EGVE '03: Proceedings of the Workshop on Virtual Environments, pp. 193–302 (2003)
11. Reiners, D., Kresse, W., 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, pp. 271–279. ACM Press, Zurich, Switzerland (2003)
12. McCrory, M., Papka, M.E., Majumder, A., Jones, D., 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)
13. Majumder, A., Stevens, R.: Perceptual photometric seamlessness in projection-based tiled displays. ACM Trans. Graph. **24**(1), 118–139 (2005)
14. Towles, H., Majumder, A., He, Z., Welch, G.: Achieving color uniformity across multi-projector displays. In : IEEE Computer Society Press (ed.) VIS '00 Proceedings of the Conference on Visualization '00, pp. 117–124. Salt Lake City, Utah, United States (2000)
15. Debevec, P.E., Malik, J.: Recovering high dynamic range radiance maps from photographs. Comput. Graph. **31**, 369–378 (1997)
16. Bern, M., Eppstein, D.: Optimized color gamuts for tiled displays. In: SCG '03: Proceedings of the Nineteenth Annual Symposium on Computational Geometry, pp. 274–281. ACM Press, San Diego, California, USA (2003)
17. Debevec, P.: Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. SIGGRAPH 98 (1998)
18. Kautz, J., Sloan, P.-P., Snyder, J.: Precomputed radiance transfer for real-time rendering in dynamic low-frequency lighting environments. SIGGRAPH (2002)
19. Lepetit, V., Fua, P.: Monocular Model-Based 3D Tracking of Rigid Objects: A survey (2005)
20. Poupyrev, I., Ikamoto, I., Kato, H., Billinghurst, M., Tachibana, K.: Virtual object manipulation on tabletop ar environment. In: International Symposium on Augmented Reality, pp. 111–119 (2000)
21. Drummond, T., Cipolla, R.: Real-time visual tracking of complex structures. IEEE Trans. Pattern Anal. Mach. Intell. **27**, 932–946 (2002)
22. Fitzgibbon, A., Simon, G., Zisserman, A.: Markerless tracking using planar structures in the scene. In: International Symposium on Mixed and Augmented Reality, pp. 120–128 (2000)
23. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. **22**, 1330–1334 (2000)
24. Stewart, C.V.: Robust parameter estimation. Soc. Ind. Appl. Math. **41**(3), 513–537 (1999)
25. Schmalstieg, D.: Designing immersive virtual reality for geometry education. IEEE Virtual Real. Alexandria, VA, USA (2006)
26. Regenbrecht, H., Wagner, M.: Interaction in a collaborative augmented reality enviromnment. In: Proceeding of Chi 2002, pp. 504–505. ACM Press, Minneapolis, Minnesota, USA (2002)
27. Shi, J., Thomasi, C.: Good features to track. In: Conference on Computer Vision and Pattern Recognition (Seattle) (1994)
28. Lowe, D.: Distincitve image features from scale-invariant keypoints. Int J Comput Vis. **20**(2), 91–110 (2001)