

IMPROVE: Advanced Displays and Interaction Techniques for Collaborative Design Review

Pedro Santos¹, André Stork¹, Thomas Gierlinger², Alain Pagani², Bruno Araújo³, Ricardo Jota³, Luis Bruno³, Joaquim Jorge³, Joao Madeiras Pereira³, Martin Witzel⁴, Giuseppe Conti⁴, Raffaele de Amicis⁴, Iñigo Barandarian⁵, Céline Paloc⁵, Maylu Hafner⁶, and Don McIntyre⁷

¹ Fraunhofer-IGD, A2,

{Pedro.Santos, Andre.Stork}@igd.fhg.de

² TU-Darmstadt, FB21,GRIS,

{Thomas.Gierlinger, Alain.Pagani}@igd.fhg.de

³ INESC-ID,

{brar, jota.costa, bruno, jaj, jap}@inesc-id.pt

⁴ GraphiTech,

{martin.witzel, giuseppe.conti, raffaele.de.amicis}@graphitech.it

⁵ VICOMTech,

{Inigo.Barandarian, Celine.Paloc}@vicomtech.es

⁶ UNIMEP,

MayluHafner@yahoo.com.br

⁷ Lighthouse,

don.mcintyre@urbanlearningspace.com

Abstract. In this paper we present evaluation results of an innovative application designed to make collaborative design review in the architectural and automotive domain more effective. Within IMPROVE, a European research project in the area of advanced displays, we are combining 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 application scenarios in the automotive domain involving FIAT Elasis from Naples, Italy and in the architectural domain involving Page/Park architects from Glasgow, Scotland. User evaluation took place at Glasgow (UK), Naples (ITA) and Darmstadt (GER), where we tested the integrated IMPROVE prototype application. The tests were conducted based on several heuristics such as ergonomics and psychomotorial factors and they were conducted based on guidelines recommended by ISO 9241 to verify whether the developed interfaces were suitable for the applications scenarios. Evaluation results show that there is a strong demand for more interactive design review systems, allowing users greater flexibility and greater choice of input and visualization modalities as well as their combination.

1 Introduction

Design Review is one of the most prominent areas benefiting from Virtual Reality and Immersive Projection Technologies. Today Virtual Reality is used to present design alternatives, but fully interactive commercial VR or AR design review applications

are still being developed or not available. Use cases generally comprise many observers in front of a back projection wall discussing a design decision on a virtual model. This is why in IMPROVE we include support for large single- or multi-tile displays as well as support for TabletPCs. In the above mentioned scenarios we can have users controlling what is visualized on the power-wall from their TabletPCs and vice versa. In addition we also aim at supporting mobile mixed reality using optical see-through HMDs which allow architects to go on site or meet around a table to look at the virtual model of a new building inserted into the real environment featuring the same lighting and reflections as the surroundings.

Components of the IMPROVE system tested by users and performance aims:

- 2D Interaction techniques as well as 3D Interaction techniques for 3D environments were assessed, as is the case when users use a TabletPC to alter the design of a model or proceed to attach annotations or choose material properties or when they use a stereo optical see-through HMD or stand in front of a Power-wall. In any case both performance and learnability were assessed.
- For the combination of display technology and rendering techniques the image quality was assessed, e.g. a comparison between standard OpenGL Lighting and GPU-based pre-computed radiance transfer.
- For the HMD ergonomic issues have been addressed such as brightness, weight, display resolution and power consumption.
- Marker-less tracking for mobile mixed reality environments has been paid special attention during the tests with respect to tracking accuracy, tracking speed, tracking area and final cost of the system.
- The mobile video transmission component has been tested to analyse whether image quality and latency times are enough for the target scenarios. Stereo video stream transmission is used to transmit high quality rendered images to a mobile user receiving the compressed 3D stereo video stream and decompressing it onto the optical stereo-see-through HMD.

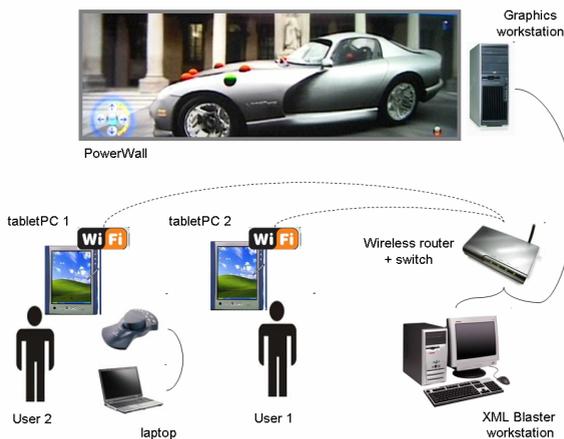


Fig. 1. A possible IMPROVE Setup

The tests of the interaction scenarios are presented in another paper in this conference. Here we present the tests of the supporting components to IMPROVE in Darmstadt on July 19th., 2006.

2 Rendering Component Usability Tests

The usability test for the rendering component was conducted to answer three questions:

- Is the rendering quality convincing?
- Is the rendering speed adequate for an interactive walk-through.

The problem with the first two questions is to find a measure for the subjective statements of being “convincing” and being “fast enough”. To answer the first question we decided to take a comparative approach. We provided the user with a rendering we judged as being convincing and asked him to compare the quality of the IMPROVE renderer to this reference image. The second question was handled by giving the user a navigation task he had to accomplish in the 3D environment. We provided 3D environments of different complexity and asked the user whether the response time of the system was good enough to accomplish the task.

2.1 Systems Setup and Features

The location of the test was the laboratory of the department A2 of IGD in Darmstadt. We were using a standard pc, namely

- CPU: Intel Pentium4 3 GHz
- Memory: 2 GB Ram
- Graphics Board: Nvidia GeForce 6600
- Operating System: Windows XP Professional Service Pack 2

To assess the rendering speed and quality we were using a desktop version of the IMPROVE renderer. The desktop version of the renderer was equipped with a simple QT user interface and supported different navigation modes, switching of lighting environments and preprocessing models. For this test we concentrated on rendering quality and speed, so the only user interaction required was navigation to evaluate the response time of the system.

The software used to create the light probe was HDRShop 1 from University of Southern California (www.hdrshop.com). This software supports the generation of high dynamic range images and manipulation of images including different selection modes and transformations between different panoramic image formats.

2.2 First Usability Test Prototype Requirements (Use Cases)

As stated in the introduction, the goal of this first user test was to validate three of the performance aims that were specified for the rendering component.

- The first one is the rendering speed
- The second goal was to validate the rendering quality.

- The third performance aim we wanted to validate is the time that is needed to capture a lighting environment by generating an HDR light probe.

2.3 Usability Test Description

The usability test of the rendering component was divided into two parts. The first one was concerned with the evaluation of the rendering engine itself, namely the rendering speed and quality. The second part was about generating new lighting input for the renderer (i.e. light probe acquisition).

2.4 Rendering Speed and Quality

The first three questions of the user test were designed to evaluate the rendering quality. The user was presented three still images (screenshots) of the same scene rendered using different renderers. The images are shown in Fig. 2..



Fig. 2. Screenshots of the same scene rendered with different quality

Image a) was generated using the second version of the IMPROVE renderer. This version supports high dynamic range reflections and direct visualization of the light probe image. Image b) was rendered using the first version of the IMPROVE renderer which supported only low dynamic range reflections and low dynamic range background images. Image c) was generated using the mental ray offline renderer. This rendering uses final gathering for environment lighting and ray tracing for calculating reflections. It should be noted, that the images a) and c) were rendered without any tone mapping while image b) has a simple tone mapping applied which is the reason for the shift in the car color. This only affects the saturation of the lacquer, not the quality of the reflections. The more washed out look of the reflections is due to the low dynamic range reflection map.

Given these three images the users were asked to rate the quality of the rendering on a scale from 1 to 5 with 5 representing the best rendering quality. The expected ordering of the images was

- Image b) worst quality of the three images
- Image a) medium quality
- Image c) highest quality

The next two questions in the user test were about the rendering speed. The users were asked to navigate the camera from a default starting point in front of a car into the position of a driver sitting in the car and looking out of the front window. The navigation was done using a mouse to drive an “examine viewer” (i.e. moving the mouse with left button pressed rotated the object about it’s center, moving the mouse

with middle button pressed pans the scene, moving the mouse with right button pressed zooms in/out). The user was asked to repeat the navigation in two scenes. The first one is the low resolution Lotus scene (see 1). This scene has 230.000 triangles (110.000 lit using PRT, 120.000 for background and glass). This scene was rendered at about 60 frames per second on the test system. The car model was downloaded from the internet (see <http://dmi.cher-alice.fr>) and was generated using some digital content creation tool (i.e. a 3D modeler, not a CAD system). The second scene showed a BMW model that was generated from CAD data (see figure 2). This model was composed of 1.35 million triangles (1.2 million lit using PRT, 150.000 for background and glass). This scene was rendered on the test system at approximately 10 frames per second.

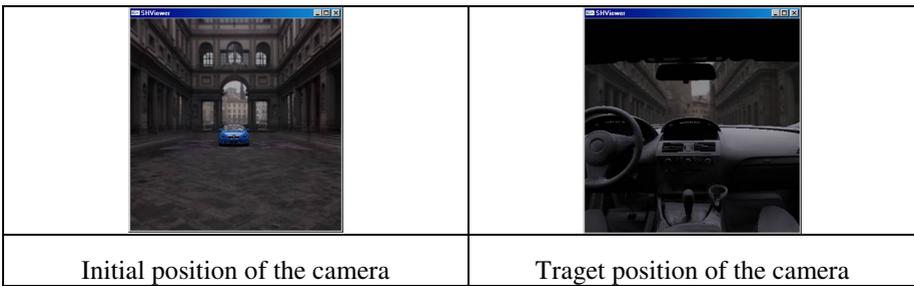


Fig. 3. BMW scene

For each of the scenes we measured the time the user required to finish the navigation and afterwards asked him whether he felt the response time of the system was fast enough to accomplish this task. The answer was given on a scale from 1 to 5 with 5 meaning the rendering was fast enough and 1 meaning the rendering was too slow.

2.5 Questionnaire Analysis

We had six test users, all with experience in navigating in 3D programs. None of the users had generated a light probe before.

Figure 4 shows the results for the first three questions about the rendering quality. The average scores show the expected result. The image generated with the first version of the IMPROVE renderer (Picture b) was rated to have the lowest quality of the series with 2.33 out of 5 points. The image generated with the second version of the IMPROVE renderer (Picture a) was rated to be of medium quality (score 3.83 out of 5) and the image generated with the offline renderer (Picture c) was rated to be the most convincing one of the series (score 4.0 out of 5). This result shows that the quality of the current IMPROVE renderer is relatively close to the quality of the offline renderer (3.83 compared to 4.0). This basically means that the real-time rendering algorithm we chose for the IMPROVE renderer (Precomputed Radiance Transfer) is adequate for rendering scenes like the one used in this test (static geometry under low frequency lighting). However, even the score of the offline renderer did not reach the maximum of 5.0 in this test. The reason for this is most likely the fact that the materials that were used on the model are relatively simple. The lacquer of the car is a Phong material, the interior

has only diffuse materials assigned. By using more advance materials / shaders it should be possible to increase rendering quality.

The results of question number four about the rendering speed are shown in figure 5. The graph on the right-hand side displays the average rating of the rendering speed in the two test scenes. It can be seen, that the rendering speed in the low polygon scene (Lotus scene) is judged to be fast enough (5.0 out of 5.0) for easy navigation. Although the scores for the complex scene (BMW scene) are considerably lower (3.17 out of 5.0) the rendering speed of this scene was still acceptable for the majority of the users. On the other hand a look at figure 6 reveals that the navigation in the complex BMW scene was harder due to the lower response time of the system. Compared to the simple Lotus scene the users needed more than twice the time to complete the navigation. This result suggests that increasing the rendering speed should not be of high priority in the next development phase. Although higher rendering speed of complex models will increase the usability of the system, the rendering quality (more specifically the quality of materials / shaders) seems to be a more pressing problem since no configuration in the rendering quality test reached the top score of 5.0 points.

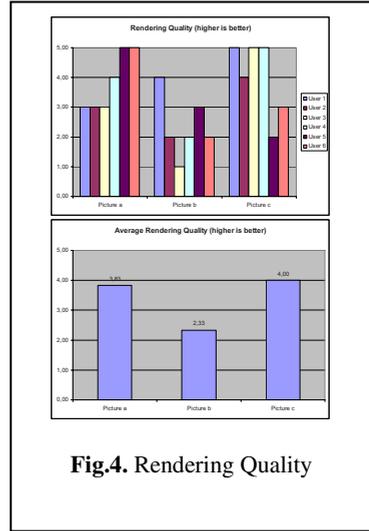


Fig.4. Rendering Quality

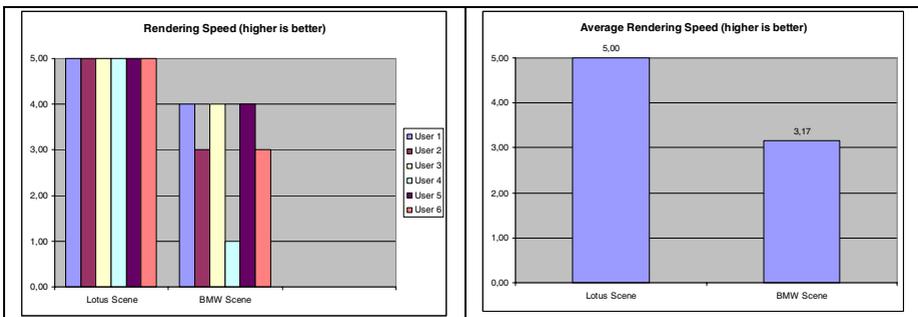


Fig. 5. Rendering Speed

2.6 Conclusion

This first evaluation of the rendering component suggests that the priority of the development should be on the improvement of the supported materials / shaders. The rendering speed is adequate for small to moderate size scenes (100.000 to 500.000 triangles). For complex scenes (around 1 million triangles) the rendering speed is acceptable, but improving the speed would increase usability. The light probe acquisition is possible in the estimated time (below one hour), but requires some training to achieve good results.

3 Video-Streaming Component Usability Tests

This section describes the results of the video-streaming component tests.

3.1 Systems Setup and Features

Since the video-streaming component is to be used to support mobile HMD users who do not have the local processing power for high-quality rendered images to be computed locally in real-time, the stereo video-streams need to be transmitted over wireless or narrow-band network connections. In addition, also the sender side can be connected to the network over narrow-band connections in worst case. Therefore we tested the video-streaming component under such circumstances. The different bandwidths were simulated by restricting the throughput at the switches the two computers were patched to. In addition to that the video-streaming sender component allows to explicitly predefine the data throughput of the stream, which was defined as follows:

- UMTS (300 Kbps video stream throughput)
- DSL (650 Kbps video stream throughput)
- WLAN (3000 Kbps video stream throughput)

The reason why we did not set the amount of video data produced to rates higher than 3000 Kbps is because the image quality would not get any better at higher values, when we compared the original stream and the encoded/decoded stream on the same computer for the resolutions we tested.

We did four tests, three with test candidates and one to measure transmission performance:

- Original Image Detection
- Comparison of video-streaming over different networks
- Responsiveness of remote interaction
- Performance measurements of the video-streaming component

For the first test, we placed two workstations next to each other and transmitted the Desktop from one machine to the other at a resolution of 1280x1024 (40% more data than 2 x SVGA) upholding 25 fps over 54Mbps WLAN to test the quality of the transmitted images with the IMPROVE system running on the sender-side.

For the second and third test the sender was set to a resolution of 1600 x 600 with a color depth of 32bit and the compressor was set to keep up a target frame-rate of 25fps on sender-side. The stream throughput to be generated was set to one of the three throughputs we wanted to test over the three different network bandwidths.

For dual channel video-streaming we use two identical threads on the sender-side which compress the images for the left eye and the images for the right eye respectively, which means left and right viewport on a horizontal span Desktop. Alternative implementations which embed both streams in one are described in D20. However since the streams are to be decompressed by two separate low-power, small form factor and low-consumption mobile computers, we split them into a dual channel stream with one channel sent to each machine. The streams are synchronized from the sender

to the two receiving sub-notebooks on the receiver-side which decompress and display the streams on the user's HMD.

In the last test we measured the performance between two workstations simulating different stream throughputs, networks and screen resolutions.

For the user tests we conducted we tested 13 individuals, 9 male and 4 female, with ages ranging between 20 - 30 years and a background in computer graphics, architecture and design.

3.2 Usability Test 1 - Original Image Detection

In the first test, the video-streaming is already taking place in 54Mbps WLAN conditions at 3000 Kbps, when the user is invited to enter the room where the test takes place. The application transmitted over the network by the video-streaming component is the current IMPROVE prototype system featuring a PRT (Pre-computed Radiance Transfer) rendered car model which remains static in its position (figure 8).

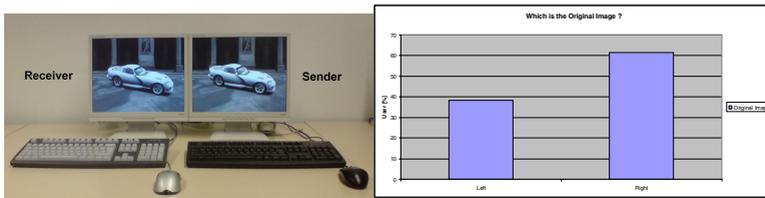


Fig. 6. Video-streaming Test Setup / Which is the original image ?

The user is then asked to find out which is the original image and which the received to see whether he can tell the difference at highest quality and almost loss-less compression. In fact the original image is the right one in Fig. 6. and differences can only be detected at very close range. The 60% of the test candidates that identified the original image correctly pointed out, that they found out it was the right image due to smooth transitions on edges which would have a sharp contrast in the original image.

As can be observed in figure 9 the outcome is pretty balanced. Although the majority of the test candidates identified the original image correctly in an average time of 14,7 seconds, the image quality appears not to have deteriorated much, because almost half of the candidates could not identify the sender correctly.

3.3 Usability Test 2 - Comparison of Video-Streaming over Different Networks

For the second test we setup the sending workstation to a resolution of 1600 x 600 horizontal span, the compressor was set up in two channel synchronized mode to uphold a framerate of 25 fps on the sender side at three different throughput rates sent over three different network simulations:

- UMTS (384 Kbps video stream)
- DSL (768 Kbps video stream)
- WLAN (3000 Kbps video stream)

The first question was on how the candidates rated the similarity between received and sent images. A car model was used as content and remained static in the scene.

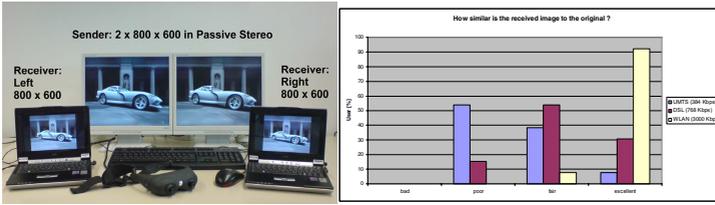


Fig. 7. User Test 2 - How similar is the received image to the original ?

As would be expected the results (Figure 9) confirm that more candidates perceive the images to be similar to each other the better the connection is. However interestingly for the static scene and for all types of connections the similarity is not rated bad by anybody. We assume that in addition to a well implemented transmission pipeline this is due to the inherent features of MPEG4 which gradually increase the quality of an image by sending delta information the longer it is not changed. Therefore most users find the similarity of the sent and received images fair to excellent in particular the longer they look at a static scene.

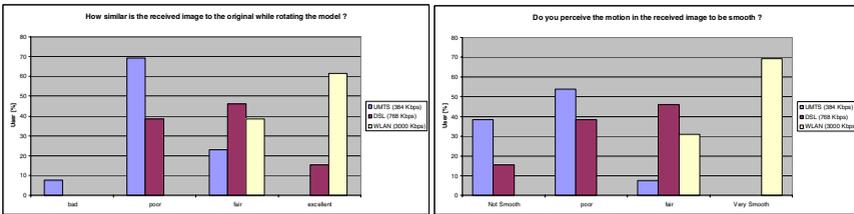


Fig. 8. User Test 2 - How similar is the received image to the original while rotating the model ? / Is the motion smooth ?

The second situation (figure 10) was to confront the candidates with moving images which cause much more traffic due to ever-changing image information. For this purpose the model of the car is rotated on the sender-side by using the keyboard tracking emulation inbuilt in the IMPROVE system. Compared to the static images we notice a slight shift in perception towards worse quality which was to be expected. However, only very few users consider the similarities to be bad even in worst bandwidth conditions. Mostly candidates complain about sharp edges turning soft and fuzzy. One candidate complained about seeing a few artifacts during transmission when we simulated UMTS bandwidth conditions.

The third question regarded motion smoothness. We asked the test candidates whether they were satisfied with the smoothness of motion in dynamic scenes over three different network types.

From the results we can conclude, that if Video-streaming is to be used at high resolutions for mobile collaborative applications such as the IMPROVE system, we need to ensure at least DSL transmission rates.

3.4 Usability Test 3 - Testing Remote Interaction Performance

In the third part of the test, the users were asked to remotely interact with a simple game, a maze. The outcome of those tests will be presented in the conference.

4 User Tests of the Supporting Hardware

We conclude that the rendering component and the video streaming component have both successfully been tested and yielded good results.

The speed of the rendering component is acceptable. The rendering quality is close to the quality of offline rendering if only scenes with static geometry and low frequency environment lighting are considered. The focus of the further development of the rendering component should be on supporting more advanced materials to increase the rendering quality. The acquisition of light probes that can be used with the rendering component is possible in less than an hour, but some experience and training is required to achieve acceptable results.

Concerning the video streaming, we have come to the conclusion that for mobile application within IMPROVE we need at least a network connection of 1 Mbps and above which is feasible since network operators start providing for higher bandwidth UMTS connections. The quality of the images transmitted is adequate for architects and automotive designers resulting in only minimal quality differences at such appropriate bandwidths. Due to the fact that we use dual channel transmission to two separate receivers we can use small form factor devices for decompression in mobile environments.

Acknowledgements. This work has partially been funded by European Commission Grant IMPROVE IST-2003-004785.

References

1. Debevec, P.E., Malik, J.: Recovering High Dynamic Range Radiance Maps from Photographs. *Computer Graphics*, In: SIGGRAPH 1997, vol. 31, pp. 369–378 (1997)
2. 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)
3. Drummond, T., Cipolla, R.: Real-Time visual tracking of complex structures. *IEEE Transaction on Pattern Analysis and Machine Intelligence* 27, 932–946 (2002)
4. 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)