DISPLAY AND RENDERING TECHNOLOGIES FOR VIRTUAL AND MIXED REALITY DESIGN REVIEW

Pedro Santos Fraunhofer-IGD, A2 <u>Pedro.Santos@igd.fhg.de</u> / www.igd.fraunhofer.de/

Thomas Gierlinger Fraunhofer-IGD, A2 Thomas.Gierlinger@igd.fhg.de / www.igd.fraunhofer.de/

André Stork Fraunhofer-IGD, A2 André.Stork@igd.fhd.de / www.igd.fraunhofer.de/

Don McIntyre The Lighthouse / Urban Learning Space don.mcintyre@urbanlearningspace.com / www.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 a variety of visualization technologies 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 user scenarios: architectural and automotive 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 targeted scenarios and focuses in particular on the rendering aspects of the project.

1. INTRODUCTION

Design Review is one of the most prominent areas benefiting from Virtual Reality and Immersive Projection Technologies. Use cases generally involve several 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 [5] in front of a projection wall. In the last couple of years Mixed Reality (MR) 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 MR 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 [4] 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 [2] explored the possibilities of mobile collaborative MR within their Studierstube system supporting various applications from scientific visualisation to interactively experiencing math and esp. geometry. With MR technology maturing, companies such as BMW [3] became curious and interested as to whether MR could 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 the automotive and architecture industries. 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.

In the architectural scenario we envisaged the use of mixed reality for the following scenarios (Figure 1). Architects want to be able to visualize their new building directly on the construction site. If possible the building should be lit according to the surrounding light conditions as to blend into reality in the best possible way. This means that we would have to develop head mounted display technology capable of being used outdoors.

An alternative scenario was the idea of being able to do design review indoors around a table and have several people using HMDs, while looking on a virtual model placed on top of the table.

In all scenarios a collaborative message passing framework should be able to allow exchange of annotations, changes and choice of display options or model parts.



Figure 1: Collaborative architectural review scenario featuring tablet PC and HMD

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 wanted to improve and extend the use of large single- or multi-tile projection technology (Figure 2).

In the automotive scenario, the idea was to have a collaborative design review session in front of a multi-tile projection wall with people sitting in front of it, each with a tabletPC of his own, being able to do annotations, individually inspect the model and contribute to the general presentation running on the main display wall.

To improve geometric and colour consistency of high resolution walls more efficient calibration techniques were required, which are implemented in Improve.



Figure 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 that will work in bright outside conditions
- 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 [1] 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

The IMPROVE system architecture consists of a collaborative VR Framework with instances communicating over a communication backbone via message passing.



Figure 3: Collaborative Design Review Indoors

IMPROVE can be used indoors and outdoors. A variety of tracking systems can be connected to the system providing the best possible pose estimation for virtual reality and mixed reality modes: Optical Marker-based tracking for indoor scenarios and optical marker-less tracking for outdoor use, which are both developed in this project.

A number of multi-modal interaction techniques have been developed for the IMPROVE VR framework, which range from speech recognition, laser-based menu interaction using gate menus to two-handed interaction for navigation and selection purposes. In addition IMPROVE supports collaborative media annotations and material properties edition (Figure 3).

In IMPROVE we develop HMD and large display technologies, which we will present in further detail in this paper.

On top of all those technologies we focus on photorealistic rendering techniques using shader technology on top of the OpenSG graphical framework, which we will describe in further detail.

5. DISPLAY TECHNOLOGIES

Powerwall: The Powerwall is a Tiled Display System (Figure 4 left). 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).

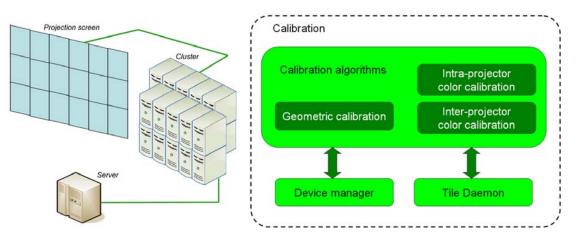


Figure 4: The Powerwall (left) and Color Calibration Software (right)

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 (Figure 4 right). 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.

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" in all HMDs done in Improve.

A second prototype has been developed featuring LCD backlighting panels with SVGA resolution and a field of view of 50 degrees, offering higher brightness and translucency. This HMD is already good enough to do indoor design review.

However one of the most important problems is transparency of the virtual models against the reality on which they are super-imposed and which will continue to shine through.

In outdoor conditions this effect is so problematic and the brightness of the displays so poor, that until today there are not many displays if any, apart from retina scanning displays, that can actually be used outside.

Therefore Improve considered a new innovative approach. Since display brightness cannot be increased to match the sun's brightness, the alternative was to build a daylight-blocking display which is also see-through and allows to block daylight pixel before superimposing virtual content (Figure 5).

Since for daylight pixel blocking we use an 8Bit greyscale mask, we are even able to cast shadows of real objects on virtual objects and vice-versa. In addition the overall effect is dramatically enhanced since we need less display brightness to still be able to see the virtual content in daylight conditions.

Due to the weight of the optics at this point of development we have decided to build a binocular type display.

To our knowledge this is one of the few daylight blocking displays existing worldwide, similar work has been done by Kiyokawa et al. [22].





Figure 5: Optical see-through HMD and Daylight Blocking Display

6. RENDERING TECHNIQUES

To answer the requirements that we identified with our industrial partners, we have developed a real-time rendering engine that supports High Dynamic Range Imaging (HDRI), Image Based Lighting (IBL), Precomputed Radiance Transfer (PRT) and Approximate Bidirectional Texture Functions (ABTFs). HDRI is used to capture real world lighting by using photographs. IBL uses the lighting information in HDRIs to render virtual models. PRT is employed to simulate dynamic soft shadows in real-time and ABTFs allow for advanced material simulation. The basis of this development is the OpenSG scenegraph system [6]. OpenSG itself is based on the OpenGL 3D graphics

library [7] and has the advantage to be a cross platform rendering system. Supported operating systems are Windows, Linux and MacOSX. Another feature of OpenSG is its support for a wide variety of display devices. It is relatively easy to render to large tiled displays as well as to different types of stereo displays. Furthermore distributed rendering in a cluster is possible. The material definitions that are available in OpenSG range from the standard OpenGL Phong and textured materials to custom shaders which may be specified using either the OpenGL shading language (GLSL)[8] or Nvidias C for graphics (Cg)[9]. The current version even supports the new programmable stage of current graphics boards (Geometry Shaders).

The following sections describe the algorithms we have implemented on top of OpenSG and give insight into the practical issues that arise when using them.

6.1 Image Based Lighting

As discussed earlier, one of the main requirements for AR rendering is to get an accurate description of the real world lighting which can then be used to render the virtual objects consistently into the real scenes. One possibility to model the real world lighting with a computer is to explicitly model all the light sources that exist in the real environment. This can be done for relatively simple scenes with only a few lights. However there are two problems with this approach: First, to accurately model the light sources we need to know its physical properties (the energy distribution). If we don't have this information we have to measure it which is time consuming and must be done for every individual light source. Second it is hard to model the more complex light sources mathematically (area light sources like windows) - the light sources used in computer graphics are usually somewhat abstract (point lights, spot lights and directional lights). We can approximate an area light source by a number of point light sources, but this raises the computational complexity during rendering: we have to calculate the contribution of every light source to the scene lighting individually which means the computational power needed increases with the number of light sources. An easier approach termed Image Base Lighting (IBL) has been proposed by Paul Debevec in his paper [10]. As the name suggests, the basic idea is to use photographs to capture incident lighting. This is done by creating an omni-directional photograph (environment map) of the real scene. Each pixel of the environment map is then associated with a direction and its value is interpreted as the amount of light that is coming from that direction. There are multiple techniques for generating such an omni-directional image:

- taking a number of images and stitching them together
- using a fish-eye lense
- taking a picture of a mirror ball
- using special spherical cameras.

The resulting image can be stored in a variety of different formats, amongst others spherical environment maps, latitude/longitude maps, and angular maps (also called light probes, see Figure 6). In the IMPROVE project we have experimented with the third and fourth method to capture omni-directional images (mirror balls and special cameras). We will give some more details on this in the next section on HDRIs.

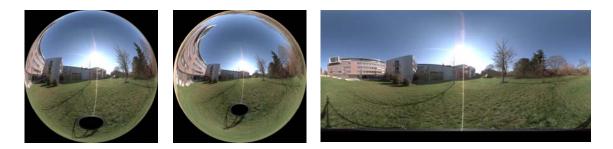


Figure 6: Environment map formats: Sphere(left), Angular(middle), Lat/Long(right)

Rendering a virtual model with an environment light is done by integrating the incident light over the hemisphere at each point on the model and weighting it according to the material properties. The main drawback of this technique

is that the light sources are assumed to be far away from the virtual model (actually they are assumed to be at infinity). Consequently local light sources cannot be simulated with this technique. However, for architectural models, where we are interested in the light incident from the sky, this restriction may be acceptable. For the automotive showroom scenario it is also true that the light sources are relatively far away from the car itself.

6.2 High Dynamic Range Imaging

Capturing the lighting information in a real scene by using photographs can be seen as using a standard digital camera as a measuring device. The problem here is that the sensors which are typically used in off the shelf cameras are not able to measure the whole dynamic range of light present in some real world scenes. Depending on the actual exposure settings of the camera one may loose either information in the dark parts of a scene (e.g. shadows appear as uniform black without any structure since part of the sensor is underexposed) or information in the very bright parts of the scene (e.g. a window through which bright sunlight falls into a room while the camera settings are adjusted to indoor lighting conditions. The window appears on the image as uniform white since the sensor is overexposed in this part). There are two possible solutions to this problem: The first one is to take multiple images of the same scene at different exposure times and to combine the structural information of the correctly exposed parts of all these images. This technique has been proposed by Paul Debevec in [11]. The second solution is to use a camera with a high dynamic range sensor. Both approaches lead to images that capture a higher dynamic range than a single photograph. Working with such images is called High Dynamic Range Imaging (HDRI). HDRI introduced the need for new image formats, since the standard formats like png or jpeg are defined only for low dynamic range information (one byte per color channel). There are now quite some formats to choose from (Radiance .hdr (or .pic)[12], ILMs OpenEXR [13], floating point tiff [14], and portable float maps (.pfm)[15]). The difference between these formats is mainly the degree of compression that is applied to the raw data which grows quit big if 32 bits are allocated for each color channel of every pixel of an HDRI.

In the IMPROVE project we use the Radiance format (which is natively supported by OpenSG). To capture HDRIs we have worked with two approaches, the use of a standard camera and a mirror ball and the use of a special HDR camera, namely the SpheroCamHDR [16]. Using a mirror ball and a standard camera has the advantage that the equipment is relatively inexpensive and it is easy and fast to acquire a low quality light probe. There are some problems with this approach that prevent it from being used to capture high quality environment HDRIs:

- The whole environment is captured with a single image which results in a rather low resolution environment map. If we use for example a 10 megapixel camera to take a photo of the sphere and store the resulting environment map in a cube map, then only less than one sixth of this resolution (about 1.6 megapixel, less than 1280x1024) is available for each side of a cube map.
- The camera that takes the picture is visible in the mirror ball. There is a technique to remove the camera from the final environment map by taking a second photo of the sphere with the viewing direction rotated by 90° (i.e. one photo from the front, another from the side). Then every part of the environment is unoccluded by the camera in at least one of the photos. The images can then be blended together to remove the camera. This technique works in theory, but in practice one has to be very accurate while placing the camera: it has to be at the same distance to the sphere for both shots and the change in viewing direction must be exactly 90°. To put it straight: we were not able to capture a light probe with this technique that did not have misalignment artefacts after blending both shots together.
- The image sequence of differently exposed images which is needed to generate a HDRI must be correctly registered the camera must not move during the captures of the different images. This is hard to achieve if the camera settings are changed manually. There are two solutions to this problem: the first one is to use computer vision techniques to register slightly misaligned images (the pfstools [17] may be used for this). The second and faster one is to make sure that the camera does not move between shots by remote controlling the camera with a computer. There is remote controlling software available for some of the midrange cameras out there. We have worked with a Canon PowerShot 640 for which the most important settings can be remotely controlled. There is a freely available software package for the Windows operating system that allows for taking HDRIs automatically by using the Canon SDK. It is called AHDRIA (Automated High Dynamic Range Image Acquisition [18]). If we ignore the reflection of the camera in the mirror ball, it is very easy and fast to capture a light probe with this setup: just place your mirror ball in your scene, setup the camera, let AHDRIA generate an

HDRI shot of the ball and then post-process the HDRI with HDRShop [19] to get a light probe image. The post-processing of the HDRI of the mirror ball is currently a manual task. We are investigating into automatizing this step. This would mean recognizing the mirror ball in the HDRI and transforming it from a sphere map into a light probe. If we had the ability to automatically capture a light probe we could update the environment light used in the architectural on-site review scenario close to real-time – lighting on the virtual model would change according to the real lighting. However – this is work in progress.

Due to the problems described above we have acquired a professional HDR camera from SpheronVR. This camera consists of a HDR column sensor that is rotated on a tripod to generate a full spherical HDRI automatically and with a very high resolution (10.600 x 5.300 pixels). The results of this device are impressive and not achievable with the mirror ball technique. On the other hand the mirror ball setup is far less expensive (the SpheroCamHDR costs about 50.000 EUR which is too much for small architectural companies).

6.3 Precomputed Radiance Transfer

Precomputed Radiance Transfer (PRT) is a technique that was introduced by Peter-Pike Sloan in 2002[20] to enable real-time rendering of some global illumination effects. Specifically it supports the simulation of soft shadows and indirect lighting for diffuse surfaces in real-time, all this under dynamic environment lighting. The idea is to precalculate for a given model how incident light is transferred into exitant radiance (light reflected to the viewer). The results of the pre-process are so called transfer functions which are projected into the function space of the Spherical Harmonics (SH). Then at run-time an actual lighting environment is also projected into the SH basis and applied to the pre-calculated transfer functions which results in a real-time rendering with the stated global effects. Mathematically the key of the technique is to move the lighting calculation into the function space of the Spherical Harmonics which reduces the integration over the hemisphere of incident light to a dot product of the coefficient vectors of the projection of transfer function and lighting environment. This technique does have some limitations: First of all the incident lighting is assumed to be infinitely far away from the model (an environment map). As stated in the section on IBL this is acceptable for our application scenarios. Second the pre-processed transfer functions are valid only for static scenes. The scene may be rotated with respect to the environment lighting but the objects of the scene are not allowed to move with respect to each other. If the scene geometry changes the pre-process has to be recalculated. However, for the exterior design review scenarios (with the user walking around a building or a car) it is acceptable. The third limitation concerns the classes of materials that are supported by the algorithm: only diffuse to glossy but not highly specular materials can be simulated. This is a rather hard limitation, especially when looking at the automotive application scenarios, since car paint is usually quite reflective. We work around this problem by separating the material properties (BRDF) into a diffuse part (which is simulated with PRT) and a highly specular part (which is simulated with environment mapping). This approach has been proposed by Green [21]. Figure 7 shows the difference between the standard OpenGL local lighting model and the global lighting effects of PRT.





Figure 7: PRT (left) vs. standard OpenGL Rendering (right)

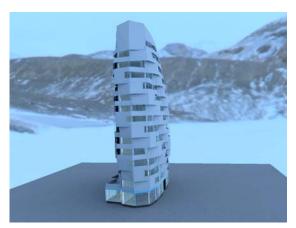
Figure 8 shows the building model in different lighting environments. Switching between lighting environments can be done in real-time using PRT – the effect of a new lighting environment is visible to user instantly.



Figure 8: The building in different lighting environments (can be changed in real-time)

6.4 Performance of the Rendering Component

In this section we give some examples and benchmarks of our rendering engine. The table below shows the speed at which the different models were rendered as well as the time it took to run the PRT pre-process. The PRT pre-process was done using 3 SH bands (9 projection coefficients) on a per vertex basis. The PRT run-time calculations were done using the GPU (the coefficients were uploaded to the graphicsboard as texture coordinates). The computer used for the benchmarks was an AMD Athlon X2 3800+ with 2 GB Ram and a Nvidia Geforce 8800 GTX graphicsboard with 768 MB of Ram. The operating system was WindowsXP with servicepack 2, the driver of the graphicsboard was Forceware 97.73.



Building Model courtesy of Page\Park architects

Triangles: 1.2 million Pre-processing time: 30 min. Speed: rendered at 100 frames per second



Lotus Model courtesy of DMI (http://dmi.chez-alice.fr)

Triangles: 110 thousand Pre-processing time: 2 min. Speed: rendered at 350 frames per second



BMW Model courtesy of BMW

Triangles: 1.2 million Pre-processing time: 30 min. Speed: rendered at 60 frames per second

7. FUTURE WORK

Until today, usage of mixed reality in outdoor environments has been a rather theoretical undertaking due to the poor performance of see-through head mounted displays. With our latest development, the daylight blocking display, we have moved a considerable step ahead towards this goal.

In upcoming projects we will establish a full HDR pipeline from image acquisition to image display, which means that we will be able to render models in their real environment as if they were real, too. Furthermore we will be able to do 3D image reconstruction to be able to gather the geometry of real objects in HDR images acquired, so we can place our virtual models in a perspective correct way into the scene and examine them, while using marker-less optical tracking to update the current viewpoint of the mobile users.

In addition to that, displays such as our newly developed daylight blocking display will help occluding real objects from virtual and vice versa and even allow for the rendering of shadows from virtual objects on real and the other way around.

We are optimistic that mobile mixed reality applications will find their way into real life applications soon.

8. ACKNOWLEDGEMENTS

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

9. REFERENCES

[1] H. Regenbrecht and M. Wagner, Interaction in a Collaborative Augmented Reality Environment, Proc. Of CHI 2002, April 20-25, 2002, Minneapolis, Minnesota, USA, pp 504-505, ACM Press 2002.

[2] D. Schmalstieg, 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

[3] G. Klinker, A.H. Dutoit, M. Bauer, J. Bayer, V Novak und D. Matzke, Fata Morgana – A Presentation System for Product Design, Intl. Symposium on Augmented and Mixed Reality, Darmstadt, ISMAR 2002.

[4] M. Fiorentino, R. de Amicis, G. Monno and A. Stork, SpaceDesign: A Mixed Reality Workspace for Aesthetic Industrial Design, Intl. Symposium on Augmented and Mixed Reality, Darmstadt, ISMAR 2002.

[5] B. Fröhlich, J. Hoffmann, K.Klüger and J. Hochstrate, Implementing Multi-Viewer Time-Sequential Stereo Displays Based on Shuttered LCD Projectors, 4th Immersive Projection Technology Workshop, Ames, Iowa, May 2004.

- [6] http://www.opensg.org
- [7] http://www.opengl.org
- [8] http://www.opengl.org/documentation/glsl/
- [9] http://developer.nvidia.com/object/cg_toolkit.html
- [10] DEBEVEC, P: Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-based Graphics with Global Illumination and High Dynamic Range Photography. *SIGGRAPH* 1998 Proceedings, ACM Press New York, USA.
- [11] DEBEVEC, P: Recovering High Dynamic Range Radiance Maps from Photographs. SIGGRAPH 1997 Proceedings, ACM Press New York, USA.
- [12] http://radsite.lbl.gov/radiance/refer/Notes/picture_format.html
- [13] http://www.openexr.com
- [14] http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf
- [15] http://netpbm.sourceforge.net/doc/pfm.html
- [16] http://www.spheronvr.de
- [17] http://www.mpi-inf.mpg.de/resources/pfstools/
- [18] http://www2.cs.uh.edu/~somalley/hdri_images.html
- [19] http://www.hdrshop.com
- [20] SLOAN, PP: Precomputed Radiance Transfer for Real-Time Rendering in Dynamic Low-Frequency Lighting Environments. *SIGGRAPH 2002 Proceedings, ACM Press New York, USA*.
- [21] GREEN, R: Spherical Harmonic Lighting: The Gritty Details. Game Developers Conference 2003, http://www.research.scea.com/gdc2003/spherical-harmonic-lighting.pdf
- [22] Kiyoshi Kiyokawa, Yoshinori Kurata, Hiroyuki Ohno, "An optical see-through display for mutual occlusion with a real-time stereovision system", Computers & Graphics, Volume 25, Issue 5, October 2001, p765-779