Photorealistic Real-Time Visualization of Cultural Heritage: A Case Study of Friedrichsburg Castle in Germany

R. Kuchar$^1$ and T. Schairer$^1$ and W. Straßer$^1$

$^1$WSI/GRIS, University of Tübingen, Germany

Abstract

This paper presents a novel highly immersive and interactive VR (virtual reality) installation targeted on photo-realistic real-time visualization. Although applicable to many other scenarios, this work is focused primarily on virtual reconstructions in the context of cultural heritage projects. We address two shortcomings in most of the current virtual reconstructions, namely interactivity and realism. On the one hand many of them are presented either as a movie or using semi-interactive techniques. In both cases the imagery is pre-rendered and therefore the visualization is lacking interactivity. On the other hand interactive real-time presentations often are neither intuitive to navigate nor visually pleasant.

We extended a real-time rendering software based on global illumination to adapt to the special needs of the visualization of virtual scenes that stem from the field of cultural heritage. A HDR (high dynamic range) daylight simulation was developed in conjunction with techniques and algorithms to significantly speed up the calculation time and increase the visual quality of the scene. To account for the different lighting situations encountered in the visualization of indoor and outdoor scenes, we developed a high dynamic range rendering pipeline that uses a dynamic tone mapping algorithm similar to human vision. To provide interactive access to the high quality 3D model even for unskilled users, we developed a very intuitive user interface based on a simple touchscreen for navigating the virtual scene. The combination of the real-time presentation of the photorealistic reconstruction and the intuitive navigation interface leads to a highly immersive and interactive VR installation. Since we are currently working on a virtual reconstruction of a Renaissance castle located in southern Germany, we will therefore use this reconstruction as a case study to present the developed features and to prove their relevance and usefulness. The virtual reconstruction is displayed using our VR installation and will be accessible to the public in the State Museum of Hohenzollern by August 2007.

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [Computer Graphics]: Interaction Techniques; I.3.7 [Computer Graphics]: Radiosity; I.3.7 [Computer Graphics]: Virtual Reality
1. Introduction

The main intention of the presented work is to give interactive access on virtual reconstructions stemming from the field of cultural heritage to the public. The task to visualize virtual reconstructed cultural heritage data, be it single artifacts, some buildings or even a whole excavation site, and presenting this work to the visitors in a museum that is open to the public is not an easy one: challenges are the type of presentation, the level of immersion the presentation should achieve and the input devices applicable for the presentation. Two of the main problems are the visualization of the virtual reconstructed data that can consist of many of millions of triangles or points) and the user interface that has to provide an easy and intuitive interaction. The easiest way to achieve this is presenting a pre-rendered movie that can be watched without any additional ways of interaction. While this can be a suitable type of presentation that has been used for years, it definitely lacks interaction and immersion. Visitors have no way to accommodate the presentation to their special interests. A widely used type of presentation can be characterized as "semi-prescribed": Visitors can interact with pre-rendered panoramic visuals, like rotating the viewing direction along the virtual horizon or zooming into and out of the pre-rendered images at special, fixed positions in the scene. The techniques, like the one presented in [LLD06], are often very similar to the Quicktime VR technology. This type of presentation is far more interactive than the above mentioned, but is still missing the immersive aspect, since for example transitions between the different points of interest are difficult to achieve. Another type of presentation is a fully interactive 3D presentation of the virtual reconstruction in a first person perspective allowing the user to navigate freely around the scene. While this is by far the most immersive and interactive type, it is also the most complicated one. Amongst the several problems that need to be addressed the most daunting ones are the creation of photorealistic scenes, the quality and rendering speed of the visualization and an intuitive interactive navigation to manipulate up to six degrees of freedom.

Therefore the most important aspect of the presented work is to display a highly realistic real-time presentation of a 3D model in a museum. Our installation offers an easy to use intuitive interface for people who are not familiar with virtual environments while providing additional information of the scene. Using our navigation interface and a reconstructed virtual model of the archaeological facility, it is a fairly easy task, for example for the archaeologist, to create a model for presentation which is based on his/her discoveries and which can be used in museums, exhibitions and for educational purposes. The work can not only be presented to a small group of scientists who have special interest in the topic, but to a much broader audience using a highly precise and scientifically serious VR model and VR installation.

In the process of doing a virtual reconstruction of the Friederichsburg Castle in Hechingen, Southern Germany, we encountered several problems, when applying standard CAD algorithms to vast indoor and outdoor scenes with lots of detail. The main challenges are lighting, rendering and navigation. The lighting of a virtual scene is one of the most important aspects of how realistic the final rendering looks. Scenes displayed using local lighting models appear very artificial and synthetic. To increase the realism and level of immersion, global illumination techniques have to be applied. The software we use for lighting calculation and visualization of the 3D scene is RadioLab: A real-time rendering software based on global illumination that implements mesh-based progressive refinement radiosity as presented in [CCWG88] to compute a highly realistic lighting of virtual scenes. We extended the system and developed a lot of enhancements, algorithms and tools in order to create an interactive, photorealistic and immersive VR experience:

Firstly, in order to simplify the creation of realistic scenes, we developed an automated realistic high dynamic range daylight simulation based on correct physical data that is used for radiosity lighting and displaying the sky dome. While radiosity lighting results in a physical correct light distribution for diffuse surfaces, care must be taken to assure that the lighting calculations can be performed in reasonable time and space. This problem gets even worse on vast reconstructed scenes that include a lot of detailed buildings that can be seen both from the exterior and interior. We applied special preprocessing techniques to keep the time and memory usage to a minimum. Secondly, since a lot of the buildings that have to be visualized in the cultural heritage sector are only illuminated from natural daylight and relatively weak artificial light sources (e.g. candles), the process of tone mapping of the radiosity values to displayable color values is not an easy task, especially when transitions from indoor to outdoor scenes can be performed by the user in real-time. We created a high dynamic range rendering pipeline that uses an automated tone mapping algorithm close to the human eye including automatic exposure over time and automatic focus. Thirdly, we added realistic visual effects like depth of field, blooming and a 3D sound system using environmental audio effects to increase the level of immersion into the scene. Lastly, we extended the RadioLab system to work hand in hand with the controller interface Commander to provide an easy to use intuitive navigation interface for people who are not familiar with virtual environments.

2. Previous and Related Work

The work presented in this paper is based upon the VR installation described in [KS03]. We extended the daylight simulation and rendering pipeline to make consistent use of high dynamic range values and significantly enhanced the quality and calculation time of the lighting simulation especially for indoor rooms. The first version of the VR installation is in use in the in-sito-reconstruction of a roman villa
in Hechingen-Stein since April 2003 and we have received a lot of feedback from users and the museum staff. We have improved the navigation interface accordingly and created a more flexible and reliable communication between the navigation interface and the visualization software.

As mentioned in the abstract and in section 1, most of the current virtual reconstructions either lack interactivity or visual quality. On the one hand, interactive presentations like the ones created by Gaitatzes et al. [GCR01] offer very interactive navigation and narrative aspects, but don’t reflect state-of-the-art lighting and rendering techniques. On the other hand, the famous “Parthenon” reconstruction by Debevec et al. [Deb05] shows a beautifully lit scene by using HDR light probes to achieve photorealistic quality. However, this work is presented in form of a movie that in turn doesn’t allow visitors to accommodate the presentation to their special interests.

We try to combine the best of these two worlds by using a photo-realistic real-time visualization and applying current lighting and rendering techniques to create an immersive and interactive VR installation.

3. The Friedrichsburg Castle

The cultural heritage site we did a virtual reconstruction of is the Friedrichsburg Castle in Hechingen, Southern Germany. This Renaissance castle had been built in 1598 by the count of Hohenzollern-Hechingen Eitelfriedrich I (1576-1605). It was destroyed in 1818 and a newer castle has been built on top of the remnants. We created the virtual reconstruction of the destroyed castle in close cooperation with Dr. Stefan Schmidt-Lawrenz, the head of the State Museum of Hohenzollern, who supervised the work guaranteeing historical correctness.

Figure 2: Detail taken from a painting of the Friedrichsburg titled “Fürstliche Jagdgesellschaft” (baronial shoot). Unknown artist, around 1730.

Only very little is known of the exact dimensions and interior of the castle. Even the single blueprint of the second level of the building, drawn by the master builder Michel d’Ixard around 1800, is lacking the relevant scale. On the one hand we had to rely on a few engravings and paintings, all subjectively distorted to represent the importance of the castle or the artists perspective; on the other hand we had access to some written contemporary descriptions and a few bills of the craftspeople that constructed the castle. Most of the measurements in this material were done in local admeasurements that varied greatly in size which in turn complicated the research on the dimensions of the castle significantly. Other contemporary renaissance castles have been regarded as a reference in both architectural style and interior design to create a very detailed and scientifically serious virtual reconstruction.

Figure 3: Abstract of a written contemporary description taken from [Rei43]

Some of the problems that we encountered during the reconstruction have been solved by using high resolution satellite imagery in conjunction with digital elevation models and detailed maps of the local surroundings. This way we could figure out for example the exact position and orientation with respect to remnants of the city wall and other renaissance buildings nearby and could rule out other locations and dimensions of the castle that were not reasonable according to the elevation model, city map or building design in this age.

4. Radiosity Lighting

The radiosity software used to generate highly realistic lighting-scenarios is RadioLab, developed at the WSI/GRIS by Ralf Sonntag in the early 90s, distributed by Pytha Lab since 1997 and improved ever since. We extended the radiosity simulation to correctly calculate the light distribution in outdoor scenes using a realistic daylight model. We are using A.J. Preethamt’s approach [PSS99] to calculate the radiance
distribution of the sky dome for a given position and day-
time. This analytical model is able to generate a realistic sun
and sky lighting in real-time. The resulting radiance values
are used in the radiosity calculation and sky dome render-
ing. The physically based lighting simulation requires a high
dynamic range representation. Displaying these radiometric
quantities directly would lead to a great loss of detail and
lighting information. A non-linear tone-mapping operator is
needed to compress the values into a displayable range. Our
virtual scene contains outdoor and indoor areas. It is nearly
impossible to compress the radiance values in a way, that the
resulting brightness is realistic for both outside and inside
parts of the model. To solve this problem, we implemented
a human-vision-oriented adaptive tone mapping technique,
which will be discussed in detail in section 5.

4.1. Adaptive Sampling

The daylight is represented by a latitude/longitude map con-
taining the radiance distribution of the sky hemisphere. We
are using an importance driven sampling technique [ODJ04]
to calculate the irradiance at every vertex of the refined scene
geometry. In a pre-processing step we generate a set of sam-
pling points on the sky map with density varying according
according to the radiance. Regions of the sky with high radiance (e.g.
the region around the sun) are sampled with a higher fre-
quency than darker areas. Every sampling point represents a
sampling direction in the scene and requires a costly visibil-
dy determination (between the sampling ray and the scene
geometry) for every vertex. The execution time of the light-
ing calculation is mainly affected by the number of sampling
points. We provide parameters to control the generation pro-
cess of the sampling point set. The user has to find a tradeoff
between a good representation of the lighting situation in the
daylight map and a reasonable number of sampling points.
Figure 4 shows a comparison between a uniform and an im-
portance sampling distribution in the upper hemisphere and
the resulting lighting quality difference. Both distributions
contain 256 sample points.

4.2. Optimization

Although the adaptive sampling scheme is a big improve-
ment compared to a uniform sampling pattern in terms of
quality and runtime, in scenes with a large amount of inte-
riors a vast number of useless sample directions is traced in
the lighting process. Only a small number of the sampling
directions “hit” the sky hemisphere through openings like
windows or doors. To improve the efficiency in these cases,
we further optimized the sampling process. The basic idea
is to split the vertices of the scene into three parts - an “out-
door”, an “indoor” and a “openings” group. This can be done
in the modelling process (by the use of layers) and does not
require additional effort.

Then we apply a two-stage technique to reduce the ini-
tial number of samples. In a pre-processing step, the sum of
the half-spaces above every opening is calculated. All sam-
ples directions, that do not point into this resulting space can
be deleted. For typical scenes, the reduction averages from
25% to 50% of the original sample set. During the light-
ing calculation, a second optimization is applied to the re-
mainning sample directions. The vertices in the outdoor group
are using the original sampling point set, while at the in-
doors vertices all sampling directions that do not intersect an
opening are rejected. The calculation of the intersection is
an easy task (vector intersecting planar rectangle) compared
to the normally used ray tracing algorithm. In scenes like
the “Rittersaal” (Figure 8(d)), using this early sample rejec-
tion mechanism, we are able to quadruple the sampling den-
sity maintaining a comparable calculation runtime. Figure 5
shows the optimization of the number of sampling points of
a given vertex located inside a room with two openings, in this case two windows next to each other in a wall.

![Figure 5](image)

**Figure 5:** Reducing the relevant sampling points of a given vertex located inside a room with two openings.

5. Rendering

When performing the rendering, we utilize the HDR representation of the scene to create state-of-the-art image based special effects. The main objective is to enhance the visual quality of the real-time presentation. Our new rendering pipeline uses HDR environment mapping, generates blooming and depth of field effects and performs dynamic time-dependent tone mapping before displaying the scene image.

All these viewpoint dependent effects are calculated using image processing methods and algorithms. Most of them are very costly considering the high resolution of today’s monitors. To guarantee real-time performance, we take advantage of the enormous computing power of current graphics hardware by applying GPGPU (General Purpose Computation on Graphics Processing Unit) techniques in our post-processing pipeline.

Originally *RadioLab* used the standard OpenGL API for rendering. The complex image processing operations require the functionality of the OpenGL 2.0 standard. For the implementation we are using FBOs (framebuffer objects) in combination with programmable shaders. FBOs serve as input (textures) and output (draw buffers) for the fragment shaders, that execute the required arithmetic operations. The data format we chose for the FBOs is the “half” data format (16-bit floating point), since the 32-bit format, that would better fit the daylighting values, is not implemented sufficiently on current graphics cards at this time. Initially the physically lit 3D-scene is rendered in to a 16-bit floating point framebuffer. At this point, also a depth texture of the current view is created.

5.1. Depth of Field

Using the actual camera geometry, aperture size, focus plane and depth values that have been transformed to world coordinates, we calculate the size of the “circle of confusion” (COC) for every pixel of the original scene image. In the second step a blur filter with a kernel size that varies according to the COC-value is applied to the scene image. This basic approach has been described in [Mit03] and further improved by Schueermann [Sch04]. We extend the visual quality of this technique by incorporating a real-time “bokeh” simulation [Mer97]. Another improvement is our image based automatic focus determination, which is indispensable in the walk-through context of the application.

5.2. Blooming

In the next stage of the pipeline we create a glare effect based on the HDR values of the input image. This phenomenon is common to all natural optical systems, like the human eye or the lens system of a photographic camera. An extensive description of the physical background is given in [SSZG95]. We use a simpler and faster approach by segmenting the image with a user-controlled brightness threshold in a first step and applying a low-pass filter afterwards. The best tradeoff between quality and performance of the blurring operation is achieved by using a separable gauss filter, implemented as two successive passes of a 1D filter kernel. The resulting glare mask is added to the result of the previous step.

5.3. Tone Mapping

This final step of the pipeline combines the intermediate images and performs a tone mapping of the resulting HDR framebuffer to a displayable 8-Bit framebuffer. The tone mapping is an essential part of the pipeline. Lighting a scene with parts of it being in- and outdoors can be far more difficult than modeling it. The brightness outside often gets exaggerated while rooms tend to be very dark, all in all the scene having unrealistically low contrast. It is impossible to find a parameter set that delivers satisfactory results using a static TM-technique. To solve this problem, we extended Reinhard’s photographic global operator [RSSF02] to respond dynamically to the amount of light in the current view as the user is moving through the virtual scene. This approach allows an automatic brightness adaption, preserving details and contrast in all areas. Every time a camera parameter changes we first measure the incident radiance, calculate the logarithmic average and apply the global operator using key- and cutoff- parameters specified by the user. We implemented the entire algorithm to run on graphics hardware. The first shader pass performs a luminance conversion and calculates the logarithmic value of every pixel in the image. Afterwards the average value is calculated utilizing hardware accelerated mip-map-filtering. For the practical use, this approach has to be enhanced in two aspects.
First we had to apply a weight function to the original image to compensate measurement discontinuities generated by areas of high radiance suddenly coming into view (e.g., a window appearing as the user is panning the point of view). Second we implemented a time-dependent adaption to new measurements, smoothing the transitions between areas of high contrast (e.g., the user is moving from indoor to outdoor part of the scene) and reducing flickering, caused by limited accuracy of the pixel-based calculations. Figure 6 shows a schematic overview of our tone mapping system.

6. VR Installation

The hardware required by our VR installation is relatively simple and low-cost. The minimal configuration is a PC with a current graphics card, a touchscreen and a display, preferably a large TFT or video projector. All interaction is performed with the touchscreen, the display is used to show a 3D view of the scene.

Using a touchscreen as the input device provides several benefits for unattended use in a museum or an exhibition: It’s not only robust and almost wear free, but it also directly connects the input and navigational display device to the user. To provide an intuitive interactive input device, two main aspects have to be considered: Location (Where am I?) and Navigation (And how do I get around here?). In everyday life persons use artificial aids, mostly maps, to ease the task of location. Virtual maps in conjunction with the current position of the virtual camera and the field of view provide a good reference for location [DS93]. To alleviate navigation in the 3D scene it is important to get the user quickly accommodated to the user interface, thus the controlling metaphor has to be straightforward [DS00].

Considering this information, we chose to display a 2D map as navigational aid and a person walking through the scene as the controlling metaphor. This metaphor implies several restrictions to the movement of the virtual camera, mainly that the movement is normally restricted to a horizontal plane. Since the 2D navigation corresponds to the 2D input device, it alleviates people with no or very little experience in navigating 3D space to understand the internal model of the interface, which on the other hand will help the user in understanding the system’s behavior from the very beginning [WO90].

6.1. Navigation in the Virtual Scene

From the software point of view, our VR installation consists of two programs that communicate via TCP/IP using various different communication APIs (Sockets, DirectPlay, ...) to ensure maximum portability across different types of underlying networks like wireless LAN, Bluetooth, etc. One of the programs, called Commander, provides the user interface, displays navigational aids and communicates with the main application RadioLab that in turn generates the 3D view of the scene. The reason for using two separate programs - one to control the movement and display valuable information such as a map and particularly highlighted points of interest, and one to display the 3D-scene - is increased flexibility and a maximized spectrum of usage. When visualizing smaller scenes, both programs can be run on one PC using a graphics card with multi-monitor support; very detailed 3D scenes with a high polygon count imply the use of two PCs, one of those with a very powerful 3D graphics card to display the 3D view.

Figure 7: The VR installation in a museum at an earlier stage used to visualize an ancient roman villa.
The Commander software displays a two dimensional top-down view of the scene along with the virtual avatar, the direction and field of view. The region on the touchscreen around the virtual avatar is subdivided into four 90 degree fields that correspond to different actions. If the user touches the region in front of the avatar, it is moving forward, the region behind the avatar is for moving backwards and the left and right region make the avatar turn accordingly. The actions associated to the four regions overlap and vary in intensity from lower intensity nearer to the neighboring regions to higher intensity in the middle of the region allowing for smooth turns and natural movement. The distance of the touching point to the avatar is also taken into account as a measure of speed of the desired action. This way the movement can be controlled very sensitive but also fast transitions to different parts of the scene are possible. Since the four regions turn according to the direction of view, the user can move around in the scene by just holding the finger pressed onto the location he wants to visit. This results in a rotation of the avatar towards the desired location, the avatar starts to walk and as it approaches the point of interest, the speed gets smoothly reduced until the location has been reached. Additionally, special points of interest can be pre-defined. These are displayed using small camera icons that can touched by the user and the avatar automatically transitions to this location. While moving around, collision detection is performed and the virtual camera is placed at a distance above ground level that corresponds to a human of average height.

A customizable menu is shown in the lower part of the touchscreen that allows the user to zoom into and out of the map, change the location to the floor above or below the current one and to switch between two modes of navigation, namely walking and examination. The walking mode has already been described above; in examination mode the user can freely look around the current position by rotating the direction of view in azimuth and elevation.

7. Results

We created a very detailed model of the castle and the surrounding city consisting of approximately 1.5 million vertices and 2 million radiosity patches. In addition to normal texture maps, bump and gloss maps were generated to reduce the amount of geometry without degrading the visual quality. This results in a total amount of 200 MB of custom textures. We applied the techniques presented in section 4 and 5 and performed radiosity lighting on the whole scene using our HDR daylight simulation. On current graphics hardware (GeForce 8800 GTX) our system achieves real-time frame rates ranging from 15 to 25 fps. In figure 8 screenshots of the resulting virtual model are shown. The VR installation itself is right now being set up in the State Museum of Hohenzollern and will be open to the public by August 2007.

8. Conclusions and Future Work

We are in close contact to the head of the State Museum of Hohenzollern and are looking forward to receive feedback.
given by the visitors. Hopefully this leads to new ideas that can be used to improve the VR installation accordingly.

Further research has to be done to simplify the inevitable preprocessing steps, namely classification and texturing: A semi-automatic classification of the openings and the spatial partitioning would reduce manual interaction considerably. By using photographs in conjunction with depth information from laser scanners or depth cameras an automated generation of bump and gloss maps could be achieved which in turn would speed up the task of texturing.

The RadioLab system is capable of displaying real-time interactive animations which can be seamlessly integrated into the VR installation. This way the scene can be enriched with animations ranging from illustrating simple mechanics to the blending of different building phases. The program Commander will be enhanced to display interactive multimedia content in the form of text, audio files and 3D animation presenting the user even more information of certain points of interest. A PDA version of the Software Commander communicating via wireless LAN, Bluetooth, etc. would be a handy tool for presentation, for example in a meeting.

To achieve an even higher level of immersion, a high-quality 3D stereo view of the scene can be provided using two video projectors and preferably wavelength multiplexed imaging [JF03], resulting in a better separation of the image pairs than using anaglyph techniques. This can be done using only one PC to generate the stereo 3D view, but since the visualization software RadioLab is also capable of synchronizing itself on two or more PCs, it is also possible to generate the left/right images on two separate machines that in turn results in higher framerates of complex scenes.

9. Acknowledgements

We would like to thank Ralf Sonntag for providing the RadioLab system and Dr. Stefan Schmidt-Lawrenz for the valuable support during the reconstruction.

References


© The Eurographics Association 2007.