

Immersive Visualization of Virtual 3D City Models and its Applications in E-Planning

Juri Engel, Hasso-Plattner-Institut, University of Potsdam, Germany

Jürgen Döllner, Hasso-Plattner-Institut, University of Potsdam, Germany

ABSTRACT

Immersive visualization offers an intuitive access to and an effective way of realizing, exploring, and analyzing virtual 3D city models, which are essential tools for effective communication and management of complex urban spatial information in e-planning. In particular, immersive visualization allows for simulating planning scenarios and to receive a close-to-reality impression by both non-expert and expert stakeholders. This contribution is concerned with the main requirements and technical concepts of a system for visualizing virtual 3D city models in large-scale, fully immersive environments. It allows stakeholders ranging from citizens to decision-makers to explore and examine the virtual 3D city model and embedded planning models “in situ”. Fully immersive environments involve a number of specific requirements for both hardware and 3D rendering including enhanced 3D rendering techniques, an immersion-aware, autonomous, and assistive 3D camera system, and a synthetic, immersion-supporting soundscape. Based on these requirements, we have implemented a prototypical visualization system that we present in this article. The characteristics of fully immersive visualization enable a number of new applications within e-planning workflows and processes, in particular, with respect to public participation, decision support, and location marketing.

Keywords: visualization system, immersive environments, virtual 3D city models, e-planning, virtual reality.

INTRODUCTION

“Major urban development projects extend over prolonged timescales, involve a large number of stakeholders, and necessitate complex decision making” (Isaacs et al., 2011). The most important decisions frequently have to be taken at an early stage of the project (Hunt et al., 2008) taking into account topological and geometrical characteristics as well as economic, social, and cultural factors (Hamilton et al., 2005). At final stages, projects typically involve external stakeholders such as the general public requiring a convincing, close-to-reality presentation of plans, variants, and processes. In both cases, fast access to urban spatial information and effective



Figure 1: Fully immersive visualization of a virtual 3D city model in the Elbe Dom facility (Fraunhofer IFF, Magdeburg, Germany).

communication tools are needed. However, today's geographic information systems (GIS) are often dominated by the view of experts and are still technically based on 2D concepts, while the underlying data has three or more dimensions (Isaacs et al., 2011).

Virtual 3D city models are essential for effective communication of complex three-dimensional urban spatial information. An increasing number of applications and systems use virtual 3D city models to integrate, manage, and visualize complex 2D and 3D urban geodata as well as associated geo-referenced thematic data (e.g., Autodesk Infrastructure Modeler, CityGRID, and CityServer3D). A growing number of cities are creating and continuing virtual 3D city models as a fundamental 3D geodata resource (Döllner et al., 2006). Meanwhile, the Open Geospatial Consortium has established the international encoding standard CityGML (Kolbe, 2009) for the representation, storage, and exchange of virtual 3D city and landscape models, implemented as an application schema of the Geography Markup Language (GML).

Virtual 3D city models are used in various application fields, e.g., driving simulations (Randt et al., 2007), disaster management (Lapierre & Cote, 2007), and e-planning (Knapp & Coors, 2008; Weber et al., 2009). Knapp and Coors (2008) developed an application for public participation via a web-browser that visualizes a virtual 3D city model to ease the understanding of planning proposals. Weber et al. (2009) use virtual 3D city models to simulate city development. Ball et al. (2007) proposed a 3D visualization system for landscape planning. While all these applications, as well as commercial tools like the Infrastructure Modeler, could be used with high resolution displays and projectors, they cannot be directly used for fully immersive virtual environments like CAVEs due to both technical and conceptual restrictions. In particular, they lack advanced visualization and rendering techniques, such as satisfactory real-time photorealistic rendering, assistive interaction techniques, and an immersion-supporting soundscape, and thus in many cases do not meet the expectations of stakeholders. Isaacs et al.

(2011) present a decision support tool that supports power walls and stereoscopic rendering. Although some advanced rendering techniques (e.g., water rendering) are implemented it still lacks other crucial techniques, e.g., ambient occlusion, dynamic sky rendering, and a 3D soundscape.

Virtual environments (VEs) represent artificial environments that are based on interactively visualizing virtual 3D models; they have manifold fields of application in Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR). As Brunnett et al. (2008) point out, “an important aspect of VR-based systems is the stimulation of the human senses – typically sight, sound, and touch – such that a user feels a sense of presence (or immersion) in the virtual environment.” For applications in e-planning, VEs require a focus on that sense-of-presence, in particular, to support not only experts but also non-experts to examine spatial structure, appearance, and relationships of plan models “in situ”. Furthermore, e-planning typically includes collaborative processes that involve a large number of heterogeneous stakeholders, whereby decisions have to be understood, discussed, and made by non-experts as well as by experts (Ross et al., 2007).

Immersive 3D virtual environments offer a complementary, intuitive access to complex 3D spatial models, broader audiences, and enable new fields of applications. Immersion is the feeling of being spatially located in the VE and thus increases its virtuality (MacEachren et al., 1999). Immersion enables better spatial understanding of complex three dimensional data (Schuchardt & Bowman, 2007), eases the creation of cognitive maps and mental frame-of-references (Pausch et al., 1997), which in turn leads to an increased performance in spatial tasks (Tan et al., 2004). Thus, immersive VEs are well-suited as tools for effective communication of complex virtual 3D city models and their complex thematic contents. Additionally, fully immersive VEs, such as CAVEs, enable exploration and analysis of 3D models by multiple users in an intuitive and collaborative way (Figure 1).

This article focuses on identifying the most important requirements and technical concepts of a system for visualizing virtual 3D city models in large-scale, fully immersive environments. We describe in detail what challenges must be faced during implementation and how to deal with them, including innovative techniques for dynamic sky rendering and semantic-based creation of an immersion-supporting soundscape. Further, we discuss our experience with the prototype presented in this article and sketch potential applications in e-planning.

HARDWARE REQUIREMENTS

To create the feeling of immersion the visualization system has to consider several technical requirements. First of all, these are hardware requirements that define an immersive environment like field-of-view, presentation scale, and display resolution.

Field-of-View

The size of the projection surface and more precisely the covered field-of-view (FOV) affects the size that an image of the VE occupies on the users retina. Any restriction of the image size on the retina decreases the immersion of the user because everything visible has an impact on the user's perception and anything else as an image of the VE reminds users that they are actually not inside the virtual 3D city model. Large displays potentially increase retinal image size and, therefore, improve immersion.

Lutz (2004) distinguish between three types of immersive VEs based on the utilized media technology and the covered field-of-view: (1) *Immersive desktop VE* cover only a certain (small) area of the FOV, (2) *semi-immersive VE* utilize large-scale displays, such as powerwalls, that cover almost the entire FOV, and (3) *fully immersive VE*. Fully immersive VEs, such as CAVEs, envelop the viewer and allow for head movement while still covering the user's FOV completely.

Presentation Scale

Another aspect that contributes to an immersive environment is the scale of the presented objects. The possible presentation scale is highly dependent on the size of the projection surface and thus on the possible field-of-view. The closer the presented size of virtual 3D objects to its size in the real world (or more exact to the expected size by the user) the higher the immersion. Ideally, in fully immersive VEs a presentation scale of 1:1 is used even in the visualization of virtual 3D city models to achieve a high degree of immersion.

Display Resolution

Given the display size, its resolution in pixels and the distance of the viewer to the display, the angular resolution (pixel size in radians) of the display can be calculated. This resolution should be high enough to create a perspicuous image, ideally as high as the maximum resolution of the human eye. If the resolution of the displays is insufficient, single pixels of the display can be distinguished. This appears unnatural and incorrect, distracts the viewer, and decreases immersion. Higher resolution buffers the case that the user move closer to the display as expected or have extraordinary sharp eyes, but also increase the computation time and thus decrease the frame rate that have to meet certain requirements, too.

3D RENDERING REQUIREMENTS

Immersive visualization implies a number of specific requirements on the real-time interactive 3D rendering system such as sufficient frame rate, creation of depth cues, intuitive interaction, and the creation of a 3D soundscape.

Frame Rate

The frame rate is the frequency at which the system displays images, mostly expressed in frames per second (FPS). Analogous to the spatial resolution of the display, the frame rate is the temporal resolution of the image stream.

The human vision system can transmit and analyze 10-12 images per second (Read & Meyer, 2000). A minimum of 15 FPS is required to create the sensation of visual continuity (Akenine-Möller et al., 2008); a frame rate of 30 FPS is usually perceived adequate. At slower frame rates changes of brightness are perceived as flicker. However, fast moving objects or camera can still create judder (non-smooth, linear motion) artifacts. These artifacts can be avoided with a higher frame rate or the usage of a motion blur filter. Any perceived judder artifacts or flicker is perceived disturbing or even cause headaches and thus reduce the immersion.

Depth Cues

Human eyes project the environment onto two-dimensional retinas. Depth perception is the ability to perceive these two dimensional images as three dimensional and thus estimate the spatial distances to and between objects. The depth sensation is generated by subconscious

interpretation of a variety of *depth cues* (Cutting & Vishton, 1995; Pfautz, 2000). These cues are also used in rendered images to enable depth perception in the artificial environment. A strong perception of depth may lead to an increased sensation of immersion (Bigoin et al., 2007).

Depth cues can be classified into monocular and binocular depending on whether they require both eyes for perception or only one, and into optical and oculomotor. Optical depth cues require interpretation of the images projected on the retina while oculomotor depth cues require interpretation of the state of the eyes muscles.

Monocular optical cues are linear perspective, relative size, texture gradient, occlusion, shadows and shading, motion parallax, areal perspective, and defocus blur. Aerial perspective is the effect of decreased contrast and saturation of distant object, as well as a color shift towards the background color, due to light scattering in the atmosphere. Defocus blur, often also referred to as depth of field, is the blurring of image parts that are not in focus. Accommodation is a monocular oculomotor depth cue and refers to the changing of the optical power of the eye to change the distance to the focal plane.

Binocular depth cues are stereopsis and convergence. Stereopsis is the impression of depth that is created by interpreting the small differences of the images of the eyes caused by their different positions. Convergence is an oculomotor depth cue referring to the convergence of the two eye balls to focus an object. Binocular depth cues are important for depth perception in near and mid-fields but have only a diminished role for objects further then 10m (Cutting & Vishton, 1995; Nagata, 1993) and thus for visualization of virtual 3D city models in 1:1 scale.

Photorealistic Appearance

The use of immersive VE for e-planning requires an accurate assessment of the specific urban situation, e.g., a planned building and its environment. Therefore, it is necessary to present the situation as accurately and realistic as possible, which requires photorealistic rendering. This implicates high requirements on the rendering quality as immersion is very fragile in a photorealistic virtual environment. Even small visual discrepancies will destroy the illusion of photorealism and with it the immersion.

Correct presentation of depth cues, in particular realistic, believable shading already creates photorealism to a certain degree. To increase the photorealism further common, every-day seen objects and phenomena, such as sky and vegetation are required. Special attention should be paid to the rendering of these objects. As most humans have a high level of visual experience with these objects their absence, unrealistic representation, or even small inconsistencies and errors are instantly noticed by the user and the scene appear unrealistic, which reduces the immersion. While there is numerous research dealing with sky and vegetation rendering (Boulanger, 2008; Jensen et al., 2001; Roden & Parberry, 2005), the implementation of the approaches in applications is still rare.

Interaction

In contrast to the passive cinematic immersion an immersive VE allows users to directly interact with the displayed virtual world and, thereby, achieves a much higher intensity of immersion. Interaction is one of the main factors that contribute to the virtuality of an artificial environment (MacEachren et al., 1999). The user should be able to interact in a natural and intuitive manner. On the one hand he should be given the full control of the six degrees-of-freedom, i.e., the user can move and rotate around the three main axes. On the other hand, unnatural or irritating

camera behavior like moving through objects as well as “getting-lost” situations should be avoided (Fuhrmann & MacEachren, 2001).

An important characteristic of an immersive visualization system is its *responsiveness*. Responsiveness is correlated with the delay with which the system corresponds to the users actions. If this delay is too long, the user loses the sense of synchrony of action and reaction, and therefore the sense of being immersed into the VE (Appino et al., 1992).

Further, the input devices should not imply restrictions on user’s physical location to allow free physical movement within the physical VE facility. The input devices should also not require the focus of the user and thus distract him from the VE. Every time users are limited in their intended actions or are distracted from the VE the sensation of immersion is reduced.

3D Soundscape

To receive a close-to-reality impression by the VE and to improve the sensation of full immersion, as much different human senses as possible should be stimulated. Assenmacher et al. (2004) consider “acoustic stimulation as a highly important necessity for enhanced immersion into virtual scenes” as our real life auditory experience is fully three-dimensional and we are capable of perceiving complex properties, like proximity, size and shape of sound sources (Blauer, 1997). A plausible 3D soundscape creates acoustic immersion and thus increases overall immersion. Paterson et al. (2010) show that a semantic-based and location-aware soundscape increases immersion and emotional engagement, which is preferable for a fully immersive VE, i.e., the soundscape should be representative and typical to the local urban sounds (Lacey & Harvey, 2011).

The 3D soundscape constitutes a core component of an immersive virtual environment. Its creation can be divided into modeling and rendering. Modeling a 3D soundscape includes searching for appropriate sound sources and proper locating them in the virtual environment, potentially attached to the associated objects. Rendering of 3D sound includes the simulation of the sound propagation considering various acoustic effects (e.g., specular reflection, diffuse reflection, diffraction, Doppler Effect, attenuation), and the final auralization, which is concerned with digital sound processing and sound output devices.

While 3D rendering libraries and systems evolved over the last decades and could establish common definitions and standard interfaces, the field of 3D sound rendering still does not have the same level of standards, systems, and libraries. Latest research in this field indicates that it is only a matter of time. Several approaches were proposed for high quality sound rendering including simulation of reflections and diffraction (Antani et al., 2012; Raghuvanshi et al., 2010; Taylor et al., 2009). Other research focuses on the ability to render a very high amount of sound sources in real-time (Moeck et al. 2007; Tsingos et al., 2004). Recent works indicate that the 3D sound rendering pipeline can be hardware-accelerated, for example, based on OpenCL (Antani et al., 2011).

While several applications require physically correct simulation of sound propagation, to increase immersion “perceptual plausibility of the auditory representation is more important than authenticity” (Blauer, 1997). In a user study of Finney & Janer (2010) their 3D soundscape for Google Streetview, despite lacking a simulation of reflections and refractions, was rated higher than real world recordings from the same location. This indicates that proper modeling of the 3D soundscape is more important for immersion than the physically correct rendering. Taylor et al. (2009) also show that for late reverberation statistical estimation is sufficient.

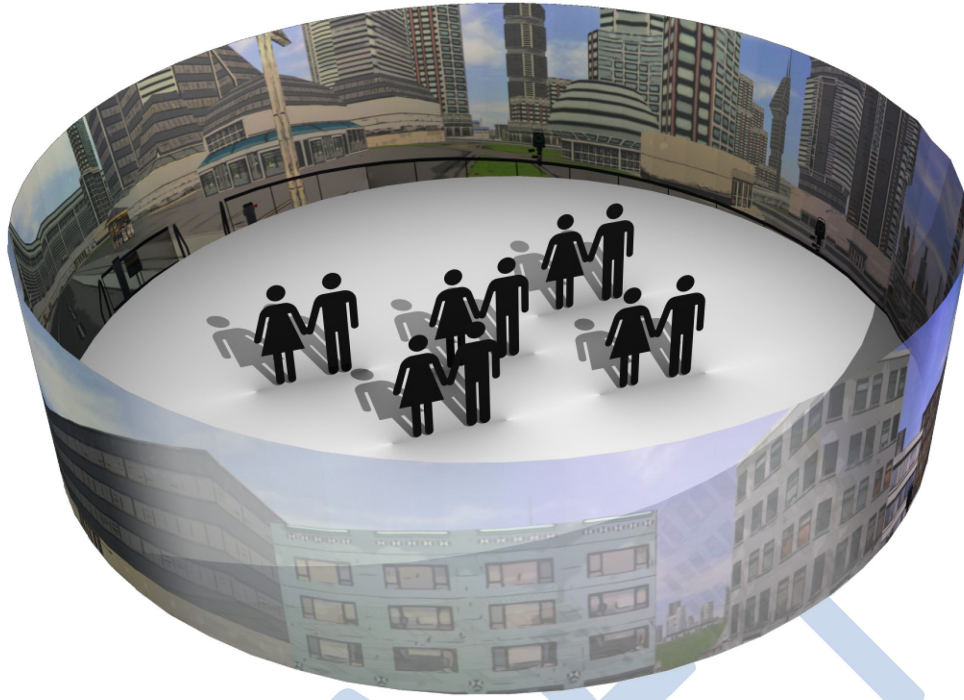


Figure 2: Conceptual illustration of the Elbe Dom facility. The viewer is completely enveloped by the 360° visualization, and the size of the projection surface enables visualization in 1:1 scale.

THE ELBE DOM FACILITY

The "Elbe Dom" facility (www.vdte.de) at the Fraunhofer IFF (Magdeburg, Germany) is a multi-user 360° cylindrical projection system suitable for large-scale interactive visualization (Figure 2). Since the opening of the facility in 2006 its application fields among others have been marketing, e-planning, and training (Belardinelli et al., 2008). In the context of e-planning it has been used to design, develop, and simulate large facilities like factories and assembly halls (Belardinelli, 2007). Previous research has shown that e-planning in form of a "digital factory" improve the efficiency and productivity of entire industrial sites (Schenk et al., 2005).

The cylindrical projection surface is 6.5m tall and has a diameter of 16m; the cylinder is bent inward at the lower part to increase the vertical field-of-view. Due to its dimensions the Elbe Dom is appropriate for visualizing large spatial 3D models, in particular virtual 3D city and building models, on a scale of 1:1. Due to its size and capacity, it especially facilitates the visualization and exploration of virtual 3D city models among a group of people.

The projection is performed using six laser projectors, each with a FOV of 68° and a resolution of 1600 × 1200 pixels, which covers 43% of the maximum resolution of the human eye (Schoor et al., 2008). The laser projectors are not capable of active or passive stereo, which has only a diminished role for creating immersion due to the dimensions and distances of the scenes visualized in the facility, anyway.

The image synthesis is performed using a render cluster with one computer (node) per projector. An additional node synchronizes the software running on the cluster and handles input devices. A tracking system with 12 infrared (IR) cameras enables determination of position and orientation of objects, e.g., controllers and users, in real-time and with a precision of 2mm. This

enables tracking of hands and fingers and, thus, wireless interaction using gestures. A touch table in the center of the user platform offers additional interaction possibilities.

The sound system, comprising of 11 loudspeakers, can be configured to create acoustic 3D scenery within an area of 4m in diameter. This enables spatialization of each sound source in the VE for multiple users. The projectors and computers are in separate rooms and acoustically imperceptible inside the projection cylinder.

SYSTEM ARCHITECTURE

A single computer is not capable to simultaneously generate images for multiple high-resolution displays/projectors in real-time. To achieve a sufficiently high frame rate, a render cluster, a compound of multiple, for rendering specialized computers (*render nodes*), is required. Ni et al. (2006) and Soares et al. (2008) provide an overview of algorithms, architectures, and technologies for high-resolution displays and cluster rendering.

Chen et al. (2001) first examined the problem of data distribution among render nodes in a render cluster. Ni et al. (2006) identified two general models: client-server and master-slave. In a client-server system the application runs on a single client that decomposes the rendering task into subtasks and delegates them with all required data to rendering servers. In a master-slave system, an instance of the application runs on every render node and stores all necessary data locally. One master node synchronizes all other nodes (slaves). A master-slave system requires usually less bandwidth, but is less transparent regarding the synchronization.

Independent of the data distribution the rendering has to be executed on multiple nodes in parallel. Parallel rendering algorithms can be classified into sort first, last, and middle (Molnar et al., 1994). The sort first algorithm segments the viewport or projection surface into tiles and assigns these to render nodes. In sort last algorithms the geometry is clustered and distributed among nodes. The generated images have to be composed to the final image afterwards, which requires image transfer between nodes. Sort middle is a combination of both: the geometry is clustered for vertex operations followed by viewport segmentation and distribution for rasterization. Image transfer and geometry clustering usually require high network bandwidth and synchronization overhead.

Several software frameworks and APIs exist, e.g., VR Juggler (Bierbaum et al., 2005), FlowVR (Allard et al., 2010), and Equalizer (Eilemann et al., 2009), that implement the algorithms described above.

In environments like CAVEs and the Elbe Dom the display, respectively the projection surface is already tiled by the projectors. This leads naturally toward a sort-first approach with each projector driven by a single cluster node. Hence, for our prototypes, we implemented a sort-first, master-slave system. This minimizes synchronization overhead, network traffic, geometric clustering, and image compositing. Every render node synthesizes images for the connected projector. The synchronization is limited to camera parameters and frame lock.

Synchronizing Render Nodes

The main issue for multi-projector systems is to ensure that the images of the different projectors appear as a single seamless high-resolution image even though it was computed on multiple render nodes. This is achieved by *gen-lock*, *data-lock*, and *frame-lock* (Soares et al., 2008). Gen-lock is the synchronization of the video signals, e.g., with shutter glasses for active stereo. It is usually implemented in hardware and requires only configuration on software side.

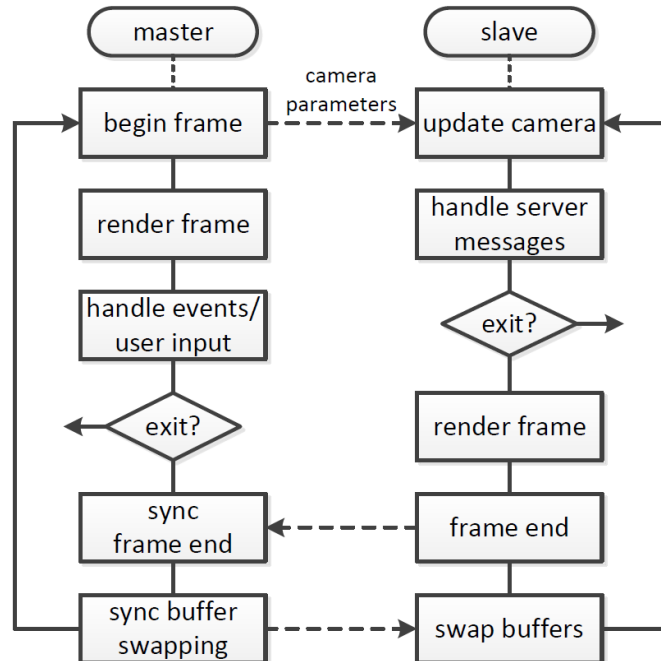


Figure 3: Overview of the synchronization process in the render cluster. At the beginning and at the end of a frame, the master node is responsible for the synchronization of the view and configuration parameters as well as a synchronous buffer swap of all slave nodes.

Data-lock is the synchronization of the data required for rendering including the camera parameters. Every change of the position and/or orientation of the virtual camera require an update on all render nodes. Figure 3 shows the synchronization concept of our implementation. At the beginning of every frame the master node broadcasts user events, virtual camera parameters and other state changes to all slave nodes, which update their local parameterization.

Frame-lock ensures that all projectors display the next frame, i.e., swap the front and back buffer, simultaneously. Missing frame-lock results in tearing-artifacts. In our implementation frame lock is achieved by synchronization at the end of every frame (Figure 3). The slave nodes inform the master that the image is rendered and ready for display and wait (frame end). The master sends a release message as soon as all slaves are finished with rendering and ready for display (sync buffer swapping).

CHALLENGES FOR VISUALIZATION SYSTEMS

As described above, to create immersion the visualization system has to meet specific requirements. The fulfillment of these requirements poses numerous challenges. In this section we describe how to deal with these challenges by the example of our visualization system.

Enhancements of the 3D Rendering Process

An important task of image synthesis in the visualization of 3D data is the creation of a sensation of depth by providing depth cues, which are described above in detail. Several of these cues are implemented within the real-time rendering pipeline (Akenine-Möller et al., 2008) and OpenSceneGraph (Wang & Qian, 2010), which is an open source 3D graphics API that we use for rendering. These depth cues are linear perspective, texture gradient, height-in-visual-field, occlusion, motion parallax and local lighting. To increase photorealism and the feeling of

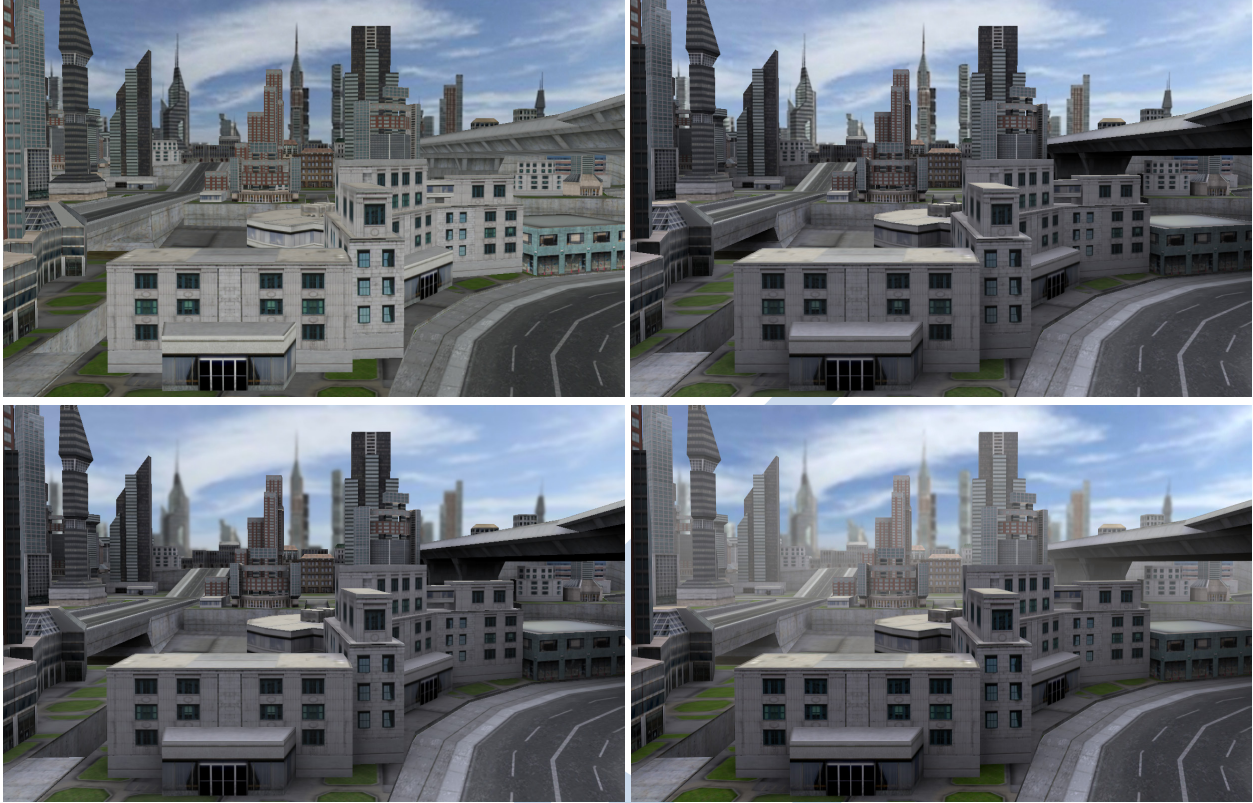


Figure 4: Depth perception can be increased by additional depth cues. Top left: original appearance; top right: shading with ambient occlusion; bottom left: ambient occlusion and background blur; bottom right: ambient occlusion, background blur and fog.

immersion we enhanced the 3D rendering process beyond the built-in functionality by implementing several extensions described in this section of the article.

Additionally to local lighting we implemented ambient occlusion (Landis, 2002) to approximate global illumination effects and thus add realism to the scene (Figure 4). Ambient occlusion takes into account the attenuation of light due to occlusion. As ambient occlusion does not depend on light direction, it can be precomputed for static objects. For night scene settings we add deferred shading (Akenine-Möller et al., 2008), an advanced rendering technique that decouples scene geometry from lighting and thus enables rendering of many light sources in real-time. We use this technique to increase the realism by adding street lamps to the urban environment (Figure 5) and thus emphasize the immersion in night and twilight sceneries.

Mather (1996) points out that “the realism of computer-generated images should also be enhanced by the addition of selective blur to background regions” because defocus blur, also known as depth-of-field, is an important depth cue. Thus, we apply blur to objects in the background. Additionally, to create the areal perspective depth cue we add fog to the scene. The effects of ambient occlusion, background blur, and fog are shown in figure 4.

Besides the implementation of the described depth cues to improve the feeling of immersion, the visualized scene should be arranged in such a way that depth cues are facilitated. For instance, in our prototype system, we add background objects, a horizon, and sky to create depth-related and scale-assigning reference objects.



Figure 5: Visualization of a virtual 3D city model at night. The photorealism is increased by using advanced rendering techniques, e.g., for rendering of high amount of light sources (street lights) and a physically correct night sky.

Dynamic, Photorealistic Sky Model

Jensen et al. (2001) describe a framework for day-night cycles, but examine most aspects of the rendering of day and night sky phenomena in isolation and provide no detailed information on intensity handling between them. For our prototype we implemented a real-time rendering technique for a physically correct dynamic sky that supports seamless transitions over the complete day-night cycle including sunrise, sunset and night sky with stars, Milky Way, and planets (Figure 5) (Müller et al., submitted). The color and brightness of the sky is calculated physically correct and in real-time for every pixel according to the atmosphere model of Bruneton and Neyret (2008). A collection of the used astronomical algorithms is given by Meeus (1994).

For star rendering our technique uses points with tiny viewpoint aligned billboards to apply point spread functions for intensity and glare. Additionally scattering and scintillations are applied using a simplified atmosphere model (Bucholtz, 1995). In contrast to a high resolution star texture (Jensen et al., 2001) this approach introduces no rendering artifacts due to sampling. Star positions are retrieved from the Yale Bright Star Catalogue (Hoffleit & Warren, 1995), which contains 9500 stars.

Jensen et al. (2001) feature photorealistic rendering of the moon using lunar surface scattering. Yapo & Cutler (2009) use photon tracing instead. Although both techniques provide photorealistic results, they are too slow for real-time visualizations. Our approach is to model the moon entirely on the GPU as a viewpoint oriented billboard. On the surface of the billboard we simulate the virtual moon sphere. This allows photorealistic rendering in real-time.

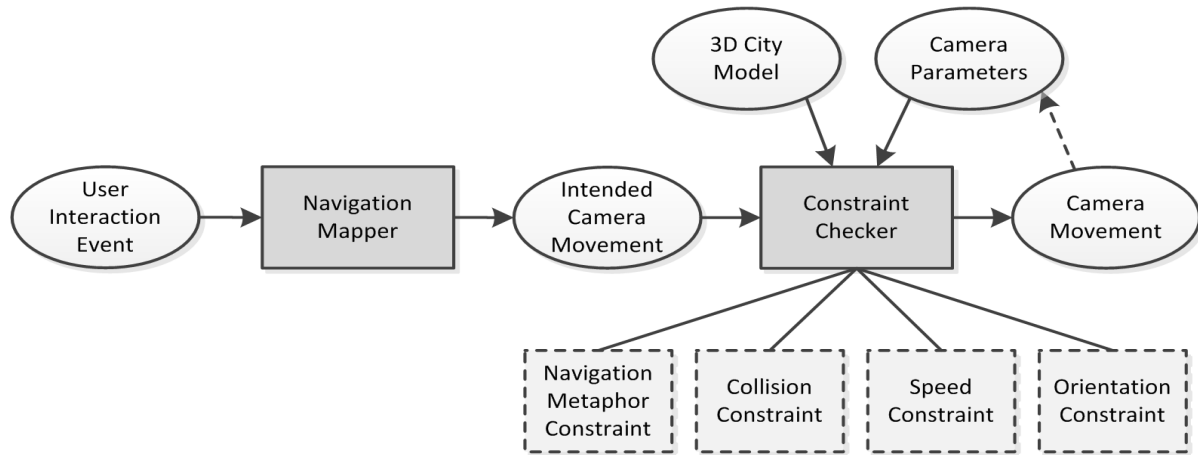


Figure 6: Immersion-aware camera system using smart interaction techniques. User input is mapped indirectly to camera movement, which is additionally checked for constraints to prevent inappropriate camera parameters.

To render clouds we use an array of precomputed 2D Perlin noise for higher cloud layers and naive scattering (Dubé, 2005) for lower layers to create a three dimensional appearance. To simulate changes over time in higher layers we use a varying density threshold.

This technique gives us the ability to create a photorealistic and physically correct sky the user is familiar with. On the one hand, this increases the immersion; on the other hand, this gives the user the ability to analyze the planning according to certain sky related properties, such as possible occlusion of the sun or appearance at a particular day- or nighttime.

For e-planning purposes, dynamic sky rendering allows us to view plan models within its surroundings under different, representatively chosen seasonal and daytime conditions. In particular, non-expert stakeholders when ask to evaluate planning proposals frequently demand for that kind of brought examination.

Immersion-Aware Camera System

As described above, interaction is essential for immersive VE. The main part of the interaction is navigation, the control of the virtual camera. Desktop input devices, such as mouse or keyboard, are not well suited for immersive VEs. They restrict user interaction: not all six DOF can be controlled simultaneously and wired devices imply restrictions on the user's physical position. Instead of desktop input devices, we utilize a wireless space mouse that enables simultaneous translation and rotation of the virtual camera. Other input devices that are suitable for immersive VEs are for example flight stick and hand tracking for gesture control.

Interaction in a fully immersive VE requires an "intuitive control". An unexpected reaction of the VE to user input reduces the acceptability of the application and thus lowers the immersion. To facilitate natural movement and to avoid collisions the user is assisted by smart interaction techniques (Buchholz et al., 2005).

Smart interaction techniques (e.g., semi-automatic pedestrian or helicopter interaction controls) indirectly map user inputs to camera movement (Figure 6) using navigation metaphors such as pedestrian or helicopter metaphor. Further, constraints for camera movement are defined to prevent collision of the virtual camera with the virtual objects and thus moving through objects and to avoid "getting-lost" situations. Such unnatural or unexpected behavior would

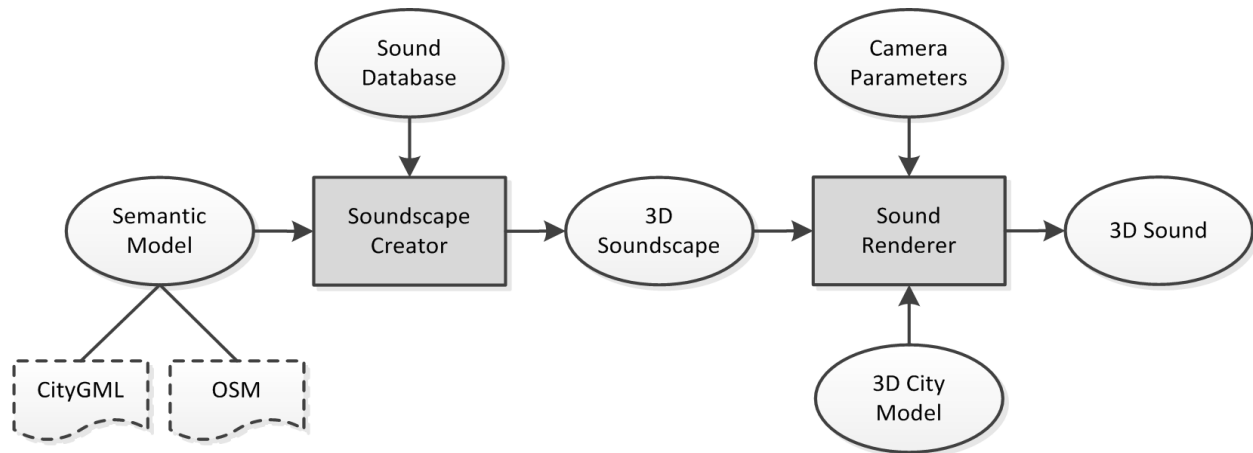


Figure 7: Creation of a plausible 3D soundscape. The samples are selected from a sound database based on semantic information and located in the virtual 3D city model. At runtime the single sound sources are synthesized to create a plausible urban sound experience.

drastically lower the feeling of immersion. For collision detection we use an image-based approach implemented on the GPU (Trindade & Raposo, 2011).

The camera system in a fully immersive environment should always be aware that the user is immersed into the VE. Immersion leads to a discrepancy between user's sense of balance and his visual sense regarding body orientation and motion. This discrepancy may lead to simulator sickness (Kolasinski, 1995). As the discrepancy increases with speed and inclination of the camera, these are capped at a certain value. Further, to prevent abrupt movement or direction changes inertia is applied to the camera control. Thus, a user-indicated change in the direction leads to an acceleration or deceleration of the camera movement in the according direction.

Immersion-Supporting 3D Soundscape

Usually applications that require a 3D soundscape concentrate on 3D sound rendering. But, we identified in the 3D rendering requirements section that modeling a perceptual plausible 3D soundscape is at least equally important for the generation of immersion. This was also our focus during implementation.

One approach is to model the 3D soundscape for a virtual 3D city model explicitly. Lacey & Harvey (2011) created a realistic 3D soundscape by recording real-world sound samples and located them in a VE. This 3D soundscape is not reusable; the process has to be repeated for every single virtual 3D city model and implies a high degree of manual work. Finney and Janer (2010) create a soundscape with samples from a free online sound database. However, these samples also need to be classified and placed manually into the VE.

Our approach is more generic (Figure 7). We model the 3D soundscape automatically by analyzing the virtual 3D city model. Ideally this is a CityGML (Kolbe, 2009) formatted model including semantic information; otherwise Open Street Map (Haklay & Weber, 2008) data is utilized to obtain this information. Based on the semantic, objects and places with typical aural appearance are identified, classified and enriched with a sound source. According to the classification the sound samples are selected from a sound database, which contains representative samples for specific urban objects, places, and situations, e.g., traffic noise, church bells, running water, and typical sounds of a pedestrian zone. Different types of sources are

supported: (1) point sources, e.g., for churches, (2) polylines, e.g., for streets, and (3) areas, e.g., for parks. For area sources only a single sample is played that serves as background sound and is representative for this source, like wind rustling in the trees of a park. Murray Schafer (1994) states that the differentiation of the soundscape into two layers (background and foreground) is more suitable for information delivery and perceptual comfort than modeling background with single sound sources (e.g., every tree in a park with its own sound).

The sound database is filled manually with sources from a free database like Freesound (www.freesound.org). Algorithms for automatic classification exist, e.g., using support vector machines (SVM), but require training of the SVM and manual post processing as the accuracy is 80% (Roma et al., 2010). We consider these algorithms for future work as it might at least provide a pre-selection of possible samples.

We select samples that were recorded in a representative environment, but are free from ambient noise, e.g., church bells from the pedestrian perspective but without traffic sound. Of course, these samples cannot perfectly match the virtual environment with regards to physically correct sound propagation, but are representative for a typical urban environment.

All information concerning the sound sources is saved in a human-readable file format. If required, this file, representing the 3D soundscape for the particular virtual 3D city model, can be further edited to modify the sound properties or to add sounds that are not supported by the soundscape creation algorithm. As the application starts this 3D soundscape file is loaded and the actual sound sources created.

At runtime the sound sources are rendered to create an acoustic immersion. For sound rendering we use the FMOD API (www.fmod.org). FMOD is capable of rendering thousands of sound sources in real time and supports sound effects like attenuation, Doppler Effect for moving sources and camera movement, and reverberation. It also takes into account the geometry of the virtual environment and factors in occlusions.

EVALUATION OF OUR PROTOTYPE

We presented our prototypical, fully immersive VE to a group of GIS experts, people of the public sector, as well as from universities at a 3D city models workshop of the German Association of Cities (Figure 1). All participants had no or only few experience with fully immersive VEs. As a case study we prepared a visualization of a synthetic city with an automatic camera path through it. Further we gave the users the opportunity to manually explore the VE using a 3D space mouse. During the study, we observed the behavior of the participants and further conducted brief, unstructured interviews.

Even though none of the participants experienced a virtual 3D city model in a fully immersive VE before, they gave very positive feedback to the visualization. They stated that they become completely immersed in the VE. During the camera path playback, we observed that users directly reacted to the camera movement, e.g., at fast sharp turns they leaned in the corner and they adapted to abrupt height changes. Afterwards, we asked participants to use the 3D space mouse to freely navigate through the VE. Although the users had only few experiences with such a device and VEs, they required almost no training to successfully navigate in the VE.

Further, they stated that they realized a plausible, convincing depth perception without active or passive stereo. Thus, it is reasonable that utilizing ambient occlusion, fog, and selective background blur in addition to depth cues created by the real-time rendering pipeline, is sufficient for fully immersive VEs. Finally, the participants stated that while they were immersed, they did not feel like a pedestrian, but a passenger of a kind of "space ship", mainly

due to the helicopter interaction control used that did not restrict the navigation to the ground level.

To achieve more quantified and comparable results we intend to perform a further user study. The research questions we intend to investigate in this study are:

- How immersive is our fully immersive VE compared to a desktop environment?
- Is there a correlation between immersion and performance in spatial tasks?

In the first part of the experiment the users will be asked to perform simple tasks like looking around, navigating to certain locations in the virtual environment, and selecting objects. This way the users should get familiar with the interaction techniques and devices. The second part will require users to perform a more complex task, like learning a route through the virtual environment and comparing alternative building models. The time required for the completion of these tasks will be measured and compared to the time of a reference group. The reference group will perform the same tasks in the same virtual 3D city model but on a desktop environment.

To answer the first research question, we will ask the participants to fill out a questionnaire. The questions will be selected from the questionnaire proposed by Witmer and Singer (1998) for the evaluation of the sense of presence in virtual environments. They can be answered by a simple rating on a seven-point scale. Witmer and Singer (1998) propose a second questionnaire (Immersive Tendencies Questionnaire) to measure the capability of the participants to get immersed. This way possible influence of individual capabilities can be considered in the evaluation.

APPLICATIONS IN E-PLANNING

Immersive visualization of virtual 3D city models, for instance within the Elbe Dom facility, offers various applications for e-planning tasks and challenges, e.g., for decision support, public presentation and participation, and marketing of projects, processes, and concepts.

The virtual 3D city model can be presented at a 1:1 scale to allow non-experts as well as experts realistic spatial evaluation, e.g., concerning visibility, lighting, vista quality, and neighborhood, by non-experts as well as by experts. The seamless 360° field-of-view generally enables full immersion. The immersion into the model space eases the understanding of complex spatial relations (Schuchardt & Bowman, 2007). To confirm that this also increases the performance in virtual 3D city models we intend to perform a user study. In urban planning projects this would ease the estimation of the effects of planned projects and help to accelerate decision making processes.

The most important benefits of fully immersive visualization with respect to e-planning include:

- “In Situ” Comparison of Models: Model and design alternatives can be compared and evaluated “in situ” of the simulated environment.
- Interactive Exploration: The interactive visualization allows stakeholders to explore plan models and their impact to their surroundings from all possible views. In particular, upcoming questions regarding specific impacts on the neighborhood can be directly checked by taking the corresponding camera view.
- Simulating Light Conditions and Moods: The support of the full day-night cycle enables the consideration and experience of a spatial location under different conditions and moods, such as sun rise, sun set, and at night as well as seasonal conditions.

- Visual Spatial Analysis: Different kinds of visual analysis tools can be applied and their results can be directly visualized within the VE. For example, lines of sight can be checked and compared.

In the last years the importance of public participation increased. The size of the Elbe Dom facility enables the interactive presentation of the virtual 3D city model to a relatively large group of citizens (30-40). The visualization in 1:1 scale emphasizes especially the communication of the virtual 3D city model from the every-day perspective. By immersion the users can experience the impact of the planned project. As a result, the understanding of the project as well as of the corresponding decisions and their consequences is increased, which can lead to greater acceptance.

Another application field is location marketing. Stakeholders can visit several possibly far separated locations “in situ” within a short period of time. Additionally, immersion creates high memorability and recognition value.

CONCLUSIONS & FUTURE WORK

Immersive 3D virtual environments are effective tools for communication of complex spatial information such as virtual 3D city models. In this article we describe requirements and concepts of a system for visualizing virtual 3D city models in large-scale, fully immersive environments and sketch its applications in e-planning, e.g., for decision support, location marketing, and public participation. Different stakeholders, from citizens to decision-makers, can explore a virtual 3D city model and examine different alternatives of an urban project “in situ”. The immersion into the model space eases the understanding of complex spatial relations and thus improves the estimation of the effects of planned projects. To enable the quantification of the benefits of full immersion we intend to perform a detailed user study.

We identified hardware and 3D rendering requirements that have to be fulfilled to achieve a high degree of immersion. We shortly introduced the Elbe Dom facility, a fully immersive environment that meets all hardware requirements and serves as an outstanding communication tool for virtual 3D city models. Further, this article discusses the conceptual and technical challenges for immersive visualization of 3D city models in the e-planning context, which we identified and managed during the development of our prototype, including specific enhancements of the 3D rendering process, immersion-aware, assistive 3D camera system, and a synthetic, immersion-supporting soundscape.

We intend to further improve our prototype by including animation, e.g., of vegetation, water, and pedestrians. This would enliven the virtual 3D city models and thus increase the sensation of immersion. Further, we are currently enhancing our prototype system to visualize models of historical places to enable virtual visits of historical sites.

REFERENCES

- Akenine-Möller, T., Haines, E., & Hoffman, N. (2008). *Real-Time Rendering 3rd Edition*. Natick, MA, USA: A. K. Peters, Ltd..
- Allard, J., Lesage, J.-D., & Raffin, B. (2010). Modularity for large virtual reality applications. *Presence: Teleoperators and Virtual Environments*, 19(2) (pp. 142–161).
- Antani, L., Galoppo, N., Lake, A., & Peleg, A. (2011). Next-Gen Sound Rendering with OpenCL. Retrieved August 9, 2012, from software.intel.com/file/39005.

- Antani, L., Chandak, A., Taylor, M., & Manocha, D. (2012). Direct-to-Indirect Acoustic Radiance Transfer, *IEEE Transactions on Visualization and Computer Graphics* (pp. 261–269).
- Appino, P. A., Lewis, J. B., Koved, L., Ling, D. T., Rabenhorst, D. A., & Codella, C. F. (1992). *An Architecture for Virtual Worlds, Presence, 1(1)* (pp. 1–17).
- Assenmacher, I., Kuhlen, T., Lentz, T., Vorländer, M. (2004). Integrating real-time binaural acoustic into VR-applications. *Virtual Environments 2004: 10th Eurographics Symposium on Virtual Environments* (pp. 129–136).
- Ball, J., Capanni, N., & Watt, S. (2007). Virtual Reality for Mutual Understanding in Landscape Planning. *International Journal of Human and Social Sciences 2(2)* (pp. 78–88).
- Belardinelli, C. (2007). Interaction of sound in virtual reality: applications to the digital factory and other uses. *Proceedings of the IV IFF, Kolloquium* (pp. 73–77).
- Belardinelli, C., Blümel, E., Müller, G., & Schenk, M. (2008). Making the virtual more real: research at the Fraunhofer IFF virtual development and training centre. *Cognitive processing 9* (pp. 217–224).
- Bierbaum, A., Hartling, P., Morillo, P., & Cruz-Neira, C. (2005). Implementing Immersive Clustering with VR Juggler. *Computational Science and Its Applications* (pp. 1119–1128).
- Bigoin, N., Porte, J., Kartiko, I., & Kavakli, M. (2007). Effects of depth cues on simulator sickness. *Proc. of the First International Conference on Immersive Telecommunications 17* (pp. 1–4).
- Blauert, J. (1997). *Spatial hearing: the psychophysics of human sound localization*. Cambridge, Massachusetts: MIT Press.
- Boulanger, K. (2008). *Real-Time Realistic Rendering of Nature Scenes with Dynamic Lighting*. Ph.D. dissertation. Orlando, FL, USA: University of Central Florida.
- Bruneton, E., & Neyret, F. (2008). Precomputed atmospheric scattering. *Computer Graphics Forum 27(4)* (pp. 1079–1086).
- Brunnett, G., Coquillart, S., & Welch, G. (2008). Virtual Realities. Dagstuhl Seminar Proceedings.
- Buchholz, H., Bohnet, J., & Döllner, J. (2005). Smart navigation strategies for virtual landscapes. *Trends in Real-time Visualization and Participation. Proc. at Anhalt University of Applied Sciences* (pp. 124–131).
- Bucholtz A. (1995). Rayleigh-scattering calculations for the terrestrial atmosphere. *Applied Optics 34* (pp. 2765–2773).
- Chen, H., Chen, Y., & Finkelstein, A. (2001). Data distribution strategies for high resolution displays. *Computers and Graphics, 25*.
- Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. *Perception of space and motion* (pp. 69–117).
- Döllner, J., Kolbe, T. H., Liecke, F., Sgouros, T., & Teichmann, K. (2006). The Virtual 3D City Model of Berlin - Managing, Integrating, and Communicating Complex Urban Information. *25th Urban Data Management Symposium (UDMS)*.

- Dubé, J.-F. (2005). Realistic cloud rendering on modern gpus. *Game Programming Gems 5* (pp. 499–505).
- Eilemann, S., Makhinya, M., & Pajarola, R. (2009). Equalizer: a scalable parallel rendering framework. *Transactions on Visualization and Computer Graphics 15* (pp. 436–452). IEEE.
- Finney, N., & Janer, J. (2010). Soundscape Generation for Virtual Environments using Community-Provided Audio Databases. *W3C Workshop: Augmented Reality on the Web*.
- Fuhrmann, S., & MacEachren, A.M. (2001). Navigation in Desktop Geovirtual Environments: Usability Assessment. *Proc. 20th ICA/ACI International Cartographic Conference*.
- Hamilton, A., Wang, H., Tanyer, A. M., Arayici, Y., Zhang, X., & Song, Y. (2005). Urban information model for city planning. *ITcon, 10* (pp. 55–67).
- Hoffleit, D., & Warren, Jr. W. H. (1995). Bright star catalogue, 5th revised ed. (hoffleit+, 1991). VizieR Online Data Catalog 5050.
- Hunt, D. V. L., Lombardi, D. R., Rogers, C. D. F. & Jefferson, I. (2008). Application of sustainability indicators in decision-making processes for urban regeneration projects. *Proc. of the ICE Engineering Sustainability, 161(1)* (pp. 77–91).
- Isaacs, J., Gilmour, D., Blackwood, D., & Falconer, R. (2011). Immersive and non immersive 3d virtual city: decision support tool for urban sustainability, *ITcon, 16*, (pp. 115–162).
- Jensen, H. W., Durand, F., Dorsey, J., Stark, M. M., Shirley, P., & Premoze, S. (2001). A physically-based night sky model. *Proc. of the 28th annual conference on Computer graphics and interactive techniques (SIGGRAPH '01)* (pp. 399–408).
- Knapp, S., & Coors, V. (2008). The use of eParticipation systems in public participation: The VEPs example. *Urban and Regional Data Management - UDMS Annual 2007*.
- Kolasinski, E. M. (1995). Simulator Sickness in Virtual Environments: Executive Summary (Tech. Rep.). U.S. Army Research Institute.
- Kolbe, T. H. (2009). Representing and Exchanging 3D City Models with CityGML. In J. Lee, & S. Zlatanova (Eds.), *3D Geo-Information Sciences* (pp. 15–31).
- Lacey, J. & Harvey, L. (2011). Sound Cartography Approaches to Urban Soundcape Research : CitySounds and Sites-of-Respite in the CBD of Melbourne. *Mapping Environmental Issues in the City* (pp. 246–265).
- Landis, H. (2002). Production-Ready Global Illumination. *RenderMan in production, SIGGRAPH 2002 Course 16* (pp. 87–102).
- Lapierre, A., & Cote, P., 2007. Using Open Web Services for urban data management: A testbed resulting from an OGC initiative for offering standard CAD/GIS/BIM services. *Urban and Regional Data Management. Proc. of the 26th UDMS*.
- Lutz, B. (2004). *Konzepte für den Einsatz von Virtueller und Erweiterter Realität zur interaktiven Wissensvermittlung*. Ph.D. dissertation, Technische Universität Darmstadt, Germany
- Mather, G. (1996). Image blur as a pictorial depth cue. *Proc. of the Royal Society B, 263(1367)* (pp. 169–172).

- MacEachren, A. M., Edsall, R., Haug, D., Baxter, R., Otto, G., Masters, R., Fuhrmann, S., & Qian, L. (1999). Virtual environments for geographic visualization: potential and challenges. *NPIVM 1999: proceedings* (pp. 35–40).
- Meeus, J. (1994). *Astronomische Algorithmen*. Leipzig, Germany: Barth.
- Moeck, T., Bonneel, N., Tsingos, N., Drettakis, G., Viaud-Delmon, I., & Alloza, D. (2007). Progressive perceptual audio rendering of complex scenes. *Proc. of the 2007 symposium on Interactive 3D graphics and games (I3D '07)*.
- Molnar, S., Cox, M., Ellsworth, D., & Fuchs, H. (1994). A Sorting Classification of Parallel Rendering. *Computer Graphics and Applications*, 14(4) (pp. 23–32).
- Müller, D., Engel, J., & Döllner, J. (2012). Single-Pass Rendering of Day and Night Sky Phenomena. *Vision, Modeling, and Visualization Workshop 2012 (VMV 2012)*.
- Murray Schafer, R. (1994). *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester, VT, USA: DestinyBooks.
- Nagata, S. (1993). How to reinforce perception of depth in single two-dimensional pictures (pp. 527–545).
- Haklay, M., & Weber, P. (2008). OpenStreetMap – User Generated Street Map. *IEEE Pervasive Computing* 7(4) (pp. 12–18).
- Ni, T., Schmidt, G. S., Staadt, O. G., Livingston, M. A., Ball, R., & May, R. (2006). A survey of large high-resolution display technologies, techniques, and applications. *Virtual Reality Conference* (pp. 223–236).
- Paterson, N., Naliuka, K., Jensen, S. K., Carrigy, T., Haahr, M., & Conway, F. (2010). Design, implementation and evaluation of audio for a location aware augmented reality game. *Proc. of the 3rd International Conference on Fun and Games*.
- Pausch, R., Proffitt, D., & Williams, G. (1997). Quantifying immersion in virtual reality. *SIGGRAPH'97 Proc. of the 24th annual conference on Computer graphics and interactive techniques* (pp. 13–18).
- Pfautz, J. D. (2000). Depth perception in computer graphics. Ph.D. dissertation, University of Cambridge.
- Raghuvanshi, N., Snyder, J., Mehra, R., Lin, M., & Govindaraju, N. (2010). Precomputed wave simulation for real-time sound propagation of dynamic sources in complex scenes. *ACM Trans. Graph.* 29(4).
- Randt, B., Bildstein, F., & Kolbe, T. H. (2007). Use of Virtual 3D Landscapes for Emergency Driver Training. In: *Proc. of the Int. Conference on Visual Simulation IMAGE 2007*.
- Read, P., & Meyer, M.P. (2000). *Restoration of motion picture film*. Oxford, U.K.: Butterworth-Heinemann.
- Roma, G., Janer J., Kersten S., Schirosa M., & Herrera P. (2010). Content-based retrieval from unstructured databases using an ecological acoustics taxonomy. *International Community for the Auditory Display*.
- Ross, L., Kleinschmit, B., Döllner, J., & Kegel, A. (2007): Geovirtual Urban Environments as Media for the Communication of Information related to Managing Urban Land. *2nd*

International Conference on Managing Urban Land - Towards More Effective And Sustainable Brownfield Revitalisation Policies.

Roden, T., & Parberry, I. (2005). Clouds and stars: efficient real-time procedural sky rendering using 3D hardware. *Proc. of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology* (pp. 434–437).

Schenk, M., Blümel, E., Straßburger, S., Hintze, A., & Sturek, R. (2004): Produktivitätssteigerung durch Virtual Reality-basierte Dienstleistungen. *Proc. des 4. Chemnitzer Produktionstechnischen Kolloquiums* (pp. 381–394).

Schoor, W., Masik, S., Mecke, R., Seiffert, U., & Schenk, M. (2008). VR Based Visualization and Exploration of Barley Grain Models with the Immersive Laser Projection System - Elbe Dom. *10th Virtual Reality International Conference* (pp. 217–224).

Schuchardt, P., & Bowman, D. A. (2007). The benefits of immersion for spatial understanding of complex underground cave systems. *Proc. of the 2007 ACM symposium on Virtual reality software and technology*.

Soares, L. P., Raffin, B., & Jorge, J. A. (2008). PC Clusters for Virtual Reality. *Int Journal of Virtual Reality*, 7(1) (pp. 67–80).

Tan, D. S., Gergle, D., Scupelli, P. G., & Pausch, R. (2004). Physically large displays improve path integration in 3D virtual navigation tasks. *Proc. of the SIGCHI conference on Human factors in computing systems*.

Taylor, M. T., Chandak, A., Antani, L., & Manocha, D. (2009). RESound: interactive sound rendering for dynamic virtual environments. *Proc. of the 17th ACM international conference on Multimedia* (pp. 271–280).

Trindade, D. R., & Raposo, A. B. (2011). Improving 3D Navigation in Multiscale Environments Using Cubemap-based Techniques. *Proc. of the 2011 ACM Symposium on Applied Computing* (pp. 1215–1221).

Tsingos, N., Gallo, E., & Drettakis, G. (2004). Perceptual audio rendering of complex virtual environments. *ACM Trans. Graph.* 23(3) (pp. 249–258).

Wang, R., & Qian, X. (2010). *OpenSceneGraph 3.0: Beginner's Guide*. Birmingham, U.K.: Packt Publishing.

Weber, B., Müller, P., Wonka, P., & Gross, M. (2009). Interactive Geometric Simulation of 4D Cities. *Computer Graphics Forum* 28(2) (pp. 481–492).

Witmer, B. G., & Singer, M. J., (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence - Teleoperators and virtual environments* 7(3) (pp. 225–240).

Yapo, T. C., & Cutler, B. (2009). Rendering lunar eclipses. *Proc. Graph. Interface* (pp. 63–69).