A Layer-based 3D Virtual Environment for Architectural Collaboration

Susanne Schmidt, Gerd Bruder, Frank Steinicke

1Department of Computer Science, University of Hamburg, Germany, {schmidt, bruder, steinicke}@informatik.uni-hamburg.de

Abstract
Architectural design processes involve a variety of users with different levels of expertise such as architects, engineers, investors or end customers. An efficient process requires all involved parties to obtain a common understanding of the architectural models and problems to be discussed. This is an ambitious task as architects as well as other involved parties often need to work with two-dimensional (2D) floor plans. While these plans are meaningful and easy to interpret for professionals, ordinary users often face problems when deducing three-dimensional (3D) properties of a building.

In this paper we address this problem by introducing an immersive virtual environment (VE) for collaborative exploration of virtual architectural models. We explore a layer-based visualization method, which stacks 2D floor plans in space providing a simple 3D impression without actually using a 3D model. Based on architectural work processes we developed a user interface including two registered representations of the same building. Our user interface allows an architect to specify a region of interest within a 3D overall view while other participants can follow his perspective in a second 2D view. In our setup the virtual building is displayed on two separate walls of an L-shaped projection system.

1. Introduction
Visualization is an important instrument in architectural review and design processes, and various computer-aided design (CAD) tools are available to support architects. Since design proposals and revisions have to be communicated between multiple people with different roles and objectives, collaboration is an essential aspect of architectural design processes. In order to get a better understanding how collaborative 3D user interfaces can support this process, we briefly summarize the typical building design process and its different stages [PS02, Tho63]:

1. Before a first architectural design can be developed, the architects have to gather information about the upcoming project. This includes interviews to identify the client’s expectations and requirements while external constraints must be considered, for example as specified by a set of regulations of the urban planning departments.
2. The analysis of the collected data provides a basis for the subsequent schematic design phase. In this stage architects develop initial ideas about basic shapes and proportions of the building as well as its appearance in the cityscape. In addition to rough sketches, physical 3D block models are used in order to facilitate a better communication and evaluation of design proposals.
3. If the conceptual proposals are approved by all decision-makers, the architects work out more details. The exact partitioning of floors into rooms is defined and interior elements are incorporated, including windows, doors, lights, power points, cable channels and so on. Furthermore, the used materials are also discussed at this stage.
4. Finally, the architects document all design decisions in a full specification, which is used as the basis for the real construction work.

Throughout all stages several iterations are required to integrate the feedback of the involved stakeholders such as building owner, investors, occupants, maintenance engineers...
or other authorities. Thus, the communication between all parties plays an essential role in successfully completing the project.

We are partners in such an urban planning process for the construction of a complex government building. Within this process, we perform focus groups and field studies with the involved parties to gain a better understanding of the requirements and constraints, and to simultaneously develop solutions, which support the entire process.

As virtual reality (VR) researchers, who have cooperated several times with architects in the past, the most surprising observation that we made right from the beginning was the limited usage of 3D visualization or VR-based exploration during these complex processes. Indeed, virtual design tools such as 3D modeling, simulation as well as game engines are becoming increasingly sophisticated, and building information modeling (BIM) is becoming a better known established collaboration process in the construction industry [Azh11]. However, many of the construction managers, architects and engineering firms still rely on and carry out most processes with 2D representations. The reasons are manifold. For instance, stakeholders usually involved in the complex construction projects such as structural engineers, electricians or other technical firms still rely on and plan their work on 2D floor plans. The plans can easily be viewed on desktops, laptops, projection walls or even printouts. Furthermore, during architectural education often two-dimensional drawings, illustrations and plans are used [Yee07]. As a matter of fact, these drafts are usually subject to many changes in the early planning stages, and adjustments in a 2D plan are less time-consuming and expensive than in their 3D counterparts.

For these reasons, 3D modeling is often outsourced to external partners. However, more and more owners are increasingly requiring BIM services from construction managers, architects and engineering firms, because non-professionals in the field of architecture often face difficulties when deducing 3D properties of a building from 2D views [BSVH10]. Miscommunication and misunderstandings at the early stages are often fatal and expensive, and hence it is crucial to identify and communicate problems and open issues in an early stage of the development process.

Putting these considerations together, in this paper we present a layer-based 3D interface prototype, which we developed in the scope of an architectural development process to allow multiple users to explore and review a building design collaboratively. The approach simply uses stacked 2D floor plans with the goal to receive a realistic impression of its 3D properties even if no actual 3D building model is available.

The remainder of this paper is structured as follows. Section 2 resumes projects that are related to the use of virtual environments (VEs) in architecture and construction. Section 3 describes the developed hardware and software setup, the visualization approach as well as concepts on how to interact with the virtual building. In Section 4 we present a pilot study that we conducted to evaluate the usage of the layer-based VE in the scope of an architectural design process that we are involved in. Section 5 concludes the paper.

2. Related Work

Architectural collaborative 3D user interfaces based on virtual reality [BP97, Why02] or augmented reality [Wan09] technologies have a variety of applications in the fields of architectural design, review, presentation, decision making and construction. The goal of most of these projects is to improve the scheduling and coordination within the professional project team and therefore save time and cost as well as supporting interactions with clients, managers and end-users [Why03].

In [PS02] several trends using VEs within the architectural domain are discussed. In particular, a lot of effort has been spent into the conceptual design and realization of VEs that aim to support architects in designing and constructing 3D buildings [AMR06, AEI03, RP’98]. The intention of those VEs is to provide users with the functionality of typical CAD tools within an immersive virtual environment (IVE) allowing a more natural and intuitive interaction.

During an architectural design process many decisions have to be made not solely regarding the entire building but also specific parts of it. In order to support these decisions, mixed reality setups have been introduced [DSD∗02, WD13], in which physical scale models are augmented with virtual parts. The users perceive this virtual content by wearing video or optical see-through head-mounted displays (HMDs).

Another approach are immersive walkthroughs, in which virtual 3D models are explored at real scale from an ego-centric perspective. For instance, [BSVH10] developed a virtual studio system for architectural design based on chroma keying approaches, in which users could perceive real-world objects in an augmented virtuality setup. In particular, users could see their own body and use tools like rulers while immersed in the VE.

Although IVES based on a tracking system and stereoscopic display can be used to explore architectural designs from an ego-centric perspective, natural interaction in such setups is limited due to restrictions of the range of tracking sensors or physical obstacles. A solution to this problem are redirected walking techniques [RKW01]. For instance, the ArchExplore project [BSH09] implemented a redirected walking strategy, which allowed users to explore a large virtual architectural model while remaining in a limited physical workspace. The project also included virtual portals to connect different virtual locations or VEs representing alternative design proposals.
Such immersive setups can provide architects and end-customers a spatial impression of a building’s room layout and interior. However, most of the existing setups only consider one user at a time. For architectural design reviews, multiple, potentially remotely distributed participants, have to be considered. The establishment of a shared interaction space enabling collaboration within such project teams is the objective of the GreenSpace project [DC96]. They provide two different interfaces: an immersive interface using a HMD and an on-screen interface. Independent of the selected interface, all participants can communicate over a network audio system. In order to support following a partner’s perspective, every user is represented by an avatar in the shared interaction space.

A comparable approach for multi-functional co-located or distributed teams is presented in [FWB13]. They describe a generalized software architecture for building a collaboration platform with different layers of abstraction. Their idea is to use two separate spaces for public team interaction and a private view respectively. In addition to a shared 3D design desktop interface, a stereoscopic real-life size digital mock-up enables users to explore the currently discussed design in an immersive way.

### 3. Layer-based 3D Virtual Environment

In this section we describe our layer-based 3D virtual environment and user interface including the hardware and software components as well as the visualization and exploration techniques.

#### 3.1. Hardware Setup

For the visualization of architectural models we use an L-Shape projection setup as the main VR environment, but we also support other VR hardware setups such as a tracked HMD real walking space. The L-Shape consists of two projection screens arranged at right angles as depicted in Figure 1. Two stereo-capable ProjectionDesign F10 AS3D projectors with SXGA+ 1400×1050 pixels resolution and 60Hz per eye update rate, one for the front and one for the floor screen, are arranged behind the L-Shape at a height of about 2 meters. In order to reduce the required distance between projectors and screens, two mirrors are used reflecting the images downward to the floor and to the back of the front screen respectively.

For stereoscopic display users wear active shutter glasses that synchronize with the projectors using DLP link technology. For head tracking five retroreflective markers are attached to the glasses. We use the ARTTRACK2 system by A.R.T., which offers high accuracy within a tracking distance of up to 4.5 meters. For good tracking coverage and robust tracking under occlusions we use seven infrared cameras, which are mounted to the ceiling and corners of the L-Shape. The tracking system is accessed by a remote workstation using the DTrack2 software. The tracked six degrees of freedom of the identified targets are sent via the VRPN protocol [VRP], which is used for 3D interaction as well as head tracking. The corresponding VRPN client software runs on an Intel graphics workstation with a Core i7 3.4GHz CPU and Nvidia Quadro K5000 graphics card. Unity3D Pro is used to render the 3D scene.

We use two input devices for interacting with the architectural 3D models. First, a Wiimote controller is connected to the workstation via Bluetooth. In the current setup only the buttons of the Wiimote are used for input. The second input device is an off-the-shelf magnifier lens that is equipped with 4 additional passive markers as illustrated in Figure 1. Its position and orientation are tracked by the ARTTRACK2 system and can be assigned to an object in the Unity3D scene.

#### 3.2. Layer Construction

As explained in the introduction, 3D data for a planned building is often not available in the early development phases. To allow users to get a spatial impression of the building even without existing 3D data, we stacked the existing 2D floor plans as illustrated in Figure 2a. Sectional views can serve as additional source for proper ceiling heights as those are essential for obtaining an intuitive understanding of proportions and dimensions.

This layer-based visualization is realized by stacking floor plans with an adjustable offset. While this step is done manually in our prototype, it can be easily automated as long as the 2D input data complies with some basic formatting rules. Besides the building itself the Unity3D scene contains a virtual representation of both the front and the floor screen. Hence, after calibration of the projectors and tracking system, the scene is an exact virtual replica of the L-Shape and therefore facilitates a proper scaling and positioning of the building.
In order to support the collaborative process, we render the 3D building model on the floor screen creating the illusion of displaying a 3D block model standing on the floor (or alternatively on a pedestal in mid-air). In this setup the front screen can display additional information like a more detailed or labeled 2D plan of the currently selected floor.

3.3. Collaborative Layer-based Interaction

The goal of this project is to create a 3D user interface for exploring architectural 3D models with a strong focus on collaboration. The described L-Shape setup is a suitable environment for this purpose as it allows multiple users to share one single interaction space. The challenge is to provide interaction concepts ensuring that all involved persons concentrate their attention on the same part of the model. To avoid the emergence of disorientation all scene navigation is executed by a single person who we refer to as the operator. This position is typically held by the architect or the person currently guiding the conversation. The operator’s head is tracked and the head’s pose is coupled to the virtual camera. Combined with shutter glasses the operator gets a realistic sense of perspective when walking around the 3D model. Other users can wear shutter glasses as well, allowing them to perceive the model stereoscopically. However, in order to get a less distorted perception of the model they stay close to the operator. In a focus group meeting with architects three different exploration modes were identified as particularly helpful. Using the buttons of the Wii mote the operator can switch between these modes:

1. In the overview mode the building is treated as a single uniform 3D object, which can be rotated around the y-axis (see Figure 2a). In this mode the overall appearance of the 3D building as well as its context can be examined without showing too much detail concerning the interior. Additional virtual contents like 2D face textures applied to the outer walls of the building or a roadmap projected onto the floor screen of the L-Shape can support this stage of exploration.

2. In the highlight mode, one particular floor can be highlighted while all overlying floors appear transparently. Additionally, all floors move upwards or downwards to guarantee the active floor is always on an easily accessible level. This mode is well suited for focussing on a specific floor without losing the general view of the building.

3. In the focus mode, the building can be folded to the currently selected floor in order to focus the user’s attention on this specific floor. As single rooms and their labels are clearly visible in this mode, it enables a more detailed exploration of the building’s interior.

The last mode provides a second 2D view of the active floor on the front screen. Thus, users can take part in the discussion of a specific floor without the need of following the operator’s movements. Furthermore, they do not have to enter the L-Shape or wear shutter glasses since the projection of the active floor plan on the front screen is displayed centered around zero parallax.

In order to guide the users’ attention on what the operator is currently focussing on we exploit a second input device, i.e., the tracked magnifier lens. The purpose of the lens is to allow the operator to specify a region of interest in the 2D floor plan. By moving the lens above the three-dimensional view of the building projected on the floor screen, a circular area is highlighted in the 2D floor plan on the front screen; all parts of the floor that are not within this region are culled out. The interaction model of the magnifier adopts the typical real-world interaction. Moving the lens in the horizontal plane results in a motion of the spot on the vertical front screen. While the 2D floor plan is magnified by an initial factor the operator can increase the zoom level by moving the lens downwards.

4. Pilot Study

In this section we describe the evaluation of the current design of the user interface in the iterative human-centered design process.

4.1. Study Design and Procedure

In order to gather information about the usability of the interface we conducted a pilot study with 5 future inhabitants (1 female, 4 males, ages 28 to 45, M=35) of a recent building planning process of the University of Hamburg, for which we received the architectural documents and, in particular, the 2D floor plans as annotated PDF documents. The building consisted of a basement as well as 11 stories as shown in Figure 2.

We performed the study with a two-stage procedure:

1. In the first stage we displayed the 2D floor plans using Adobe Acrobat on a 55 inch multi-touch tabletop, around which the participants were gathered.

2. The second stage consisted of the participants moving over to the L-Shape projection setup, in which the floors were displayed using the layer-based VE described in Section 3.

We purposely chose a simple PDF viewer to receive an impression of our prototype’s features and interaction techniques rather than comparing our system to a professional CAD software with a specific toolset. The tasks we gave the participants consisted of finding their future office rooms in the building, and following the paths they had to take to reach the rooms from the main entry of the building. We asked the participants to discuss the observations they made about the spatial properties of the building during the collaborative process using the think-aloud protocol [Lew82]. While the task involved mainly collaborative touch interaction in the tabletop setup, one participant...
mean task load of 67.59 (SD) NASA TLX questionnaire show a significant difference for Wilcoxon signed ranks tests [Ott15]. The results of the We analyzed the questionnaire data with Task Load Index (TLX) questionnaires [Har06] and performed a de-

4.2. Results and Discussion

In the following we present the questionnaire results and subjective comments during the pilot study.

Task Load We analyzed the questionnaire data with Wilcoxon signed ranks tests [Ott15]. The results of the NASA TLX questionnaire show a significant difference for mean task load of 67.59 (SD = 21.29) for the table condition and 26.97 (SD = 21.69) for the L-Shape condition, $Z = 2.02, p = .04$. In particular, we found a significant difference for mental demand between the table ($M = 68.60, SD = 23.01$) and L-Shape ($M = 29.00, SD = 27.63$) conditions, $Z = 2.02, p = .04$. We found no significant difference for physical demand between the table ($M = 35.20, SD = 28.35$) and L-Shape ($M = 24.40, SD = 25.16$) conditions, $Z = .944, p = .35$. We found a significant difference for temporal demand between the table ($M = 59.20, SD = 24.77$) and L-Shape ($M = 23.20, SD = 18.57$) conditions, $Z = 2.02, p = .04$. Also, we found a significant difference for performance between the table ($M = 70.80, SD = 30.34$) and L-Shape ($M = 24.40, SD = 19.26$) conditions, $Z = 2.02, p = .04$. Moreover, we found a significant difference for effort between the table ($M = 69.00, SD = 15.54$) and L-Shape ($M = 28.40, SD = 23.27$) conditions, $Z = 2.02, p = .04$. Additionally, we found a significant difference for frustration between the table ($M = 72.20, SD = 26.25$) and L-Shape ($M = 21.60, SD = 21.30$) conditions, $Z = 2.02, p = .04$. The results indicate that completing the task in the L-Shape was significantly less demanding than interpreting the floor plans when they were displayed on the table.

Subjective Comments We grouped the comments during the think-aloud and debriefing sessions, and identified four main topics:

1. Misinterpretations: Throughout the exploration process, a number of uncertainties regarding the architectural annotations occurred. In the layered setup, most of these uncertainties, e.g., concerning the role of a room, could be resolved by the participants by magnifying the region of interest. However, the zooming tool in the PDF was rarely used. The participants also approved the cut-outs in the layered visualization as they helped them to identify wall penetrations. A recurrent point of discussion in the PDF visualization was the currently selected floor, since it did not match the page number. Besides these differences, there were also some annotations in the 2D floor plans that could not be interpreted without help by an architect, e.g., the markings that indicated the direction of stairs. In the layered view these questions could partially be resolved due to contextual information or switching back and forth between different layers.

2. Navigation: The process of navigating through the building and finding the entrance took much longer in the PDF than in the layered visualization. Besides the learning effect that appears in the second phase of the experiment, this can be reasoned by the just mentioned misinterpretations regarding the current floor level. The navigation was further aggravated by the fact that all scrolling and zooming operations in the PDF caused a sequential reload of the page content. Additionally, multiple participants stated that switching and selecting floors in the L-Shape is much easier and faster than in the PDF and therefore improves the navigation through the building.

3. Sense of space: After every phase, we asked the participants to show the actual location of a specific room in a physical 3D model of the building. In both versions the participants were able to point to this location correctly. However, the ceiling height of a floor could not be inferred from the PDF plans, which was criticized by one participant in the debriefing. In general, the test group

![Photos taken of the three different visualization modes.](image-url)
stated that the L-Shape setup provided a better spatial impression of the building than the PDF.

4. **Collaboration:** Regarding the collaborative aspect of the two compared user interfaces, the opinions in our test group were divided. While some participants felt that it is easier to directly point on a specific location in the PDF, others preferred the magnifier interface for showing something to their group partners. The designation of an operator in the L-Shape setup was judged favorably as it prevents conflicting user inputs.

In conclusion, one participant remarked, that he would prefer to have an actual physical 3D model compared to both visualizations. However, in the absence of such a physical model the layered virtual view was preferred over the PDF.

5. **Conclusion**

In this paper we have presented a layer-based 3D VE which allows the collaborative exploration of building designs in the context of architectural review and design processes. We proposed a stacked layout based on 2D floor plans to facilitate the discussion and evaluation of designs in early development stages when 3D models are not available yet. The stereoscopic 3D view of the building was supplemented by a monoscopic representation to support interactions with multiple users. To attain a more natural interaction we also introduced different input devices such as a magnifier lens.

In the future it is important to evaluate the comfort and effectiveness of the proposed interaction concepts. For that purpose, a more extensive study involving other existing CAD tools can be conducted. A further comparison with a fully-fledged 3D model would also be useful to investigate which spatial characteristics can be explored in the layer-based visualization and which cannot. We also intend to study the effects of lifting the projection surface to a higher level, for instance by placing a purpose-built table on top of the L-Shape’s floor panel. Aside from reducing the accommodation-convergence conflict, this setup allows to project the active floor onto the table with zero parallax, which means it is displayed perspectively correct for all users. Finally, we will expand our work concerning hybrid shared interaction spaces.

**References**


[AMR06] **Aliaksyeu D., Martens J.-B., Rautenberg M.**: A computer support tool for the early stages of architectural design. *Interacting with Computers* 18, 4 (2006), 528–555. 2


