Eurographics Symposium on Virtual Environments (2004)
S. Coquillart, M. Göbel (Editors)

Image-based Techniques for the Creation and Display of Photorealistic Interactive Virtual Environments

George Drettakis[1 †], Maria Roussou[2 ‡], Nicolas Tsingos[1], Alex Reche[1] & Emmanuel Gallo[1]

[1] REVES/INRIA Sophia-Antipolis, France, [2] University College London, UK

Abstract

In this paper we introduce an image-based 3D capture process for the creation and display of photorealistic virtual environments (VEs). The resulting VEs aim to realistically recreate existing real-world scenes that can be displayed in a range of immersive VR systems using a high-quality, view-dependent algorithm and further enhanced using advanced vegetation, shadow display algorithms and 3D sound. The scenes, an archaeological site and an urban environment, were chosen according to real-world applications in the areas of urban planning/architecture and cultural heritage education. The users in each case are able to reconstruct or manipulate elements of the VEs according to their needs, as these have been specified through a detailed user requirements survey. Furthermore, a user task analysis and scenario-based approach has been adopted for the design of the virtual prototypes and the evaluation, which is currently underway. This work is being developed in the context of the EU-funded research project CREATE and the first examples of the prototype system in use are described and demonstrated in this paper.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [COMPUTER GRAPHICS]: Three-Dimensional Graphics and Realism; I.4.1 [IMAGE PROCESSING AND COMPUTER VISION]: Digitization and Image Capture.

1. Introduction

The creation of realistic 3D models has, typically, been a time-consuming labour-intensive manual process, requiring well-trained artists. Recent advances in computer graphics and vision [DTM96] allow the creation of realistic models based on photographs that are of very high-quality. Many methods [GGSC96, BBM*01] have been developed to perfect and display such models. These methods, however, have been specifically developed for conventional computer graphics rendering and not for real-time display in immersive virtual reality systems. Even if the integration of realistic models in a real-time VR environment is achieved, the resulting captured environment looks static while the display of additional elements such as shadows or vegetation is particularly difficult. Moreover, apart from the visual effect, the overall perception of realism can be further enhanced by the addition of 3D sound; however, VEs have become more and more complex recently, and real-time audio simulation requires novel algorithmic solutions to allow its integration.

The emphasis on enhancing realism in a virtual environment, due to the extended effort and performance required to achieve it, leaves little room to the development of meaningful interactivity. Highly interactive virtual environments, such as virtual prototyping systems, are usually comprised of simplified 3D models that are optimised as to not consume too many graphics rendering cycles. Hence, the interactive manipulation of photorealistic elements has remained an elusive goal.

The design of realistic and highly interactive virtual environments is a challenging task since the combination of realism and interactivity adds extra difficulty to development and implementation. We believe that both photorealism and interactivity are essential for a number of VR applications. We have therefore chosen to address this challenge through

† http://www.inria.fr/reves, George.Drettakis@inria.fr
‡ Also affiliated with the Foundation of the Hellenic World, Greece

an iterative, user-centered approach that carefully takes into account the application and user needs.

In this paper we present our approach and first solutions to this issue, focusing our attention on two areas, realism (both visual and auditory) and interactivity. For realism, we introduce a novel workflow for capture and realistic display of existing real environments based on photographs. By introducing view-dependent display techniques we achieve unprecedented realism for VEs. We further enhance these environments with efficient display of realistic vegetation, dynamic interactive shadows and realistic spatialised 3D sound. A number of design choices had to be made to allow these algorithms to integrate well in the end-user immersive VR systems. A first prototype of the system has been implemented for different immersive displays (an immersive workbench, a curved-screen and a cubic immersive display) and demonstrated in the accompanying video. Further prototypes, which are currently underway, will integrate the many interactive features that have been specified by the users and will then be evaluated on-site. The above work is being developed in the context of the EU-funded IST research project CREATE[†] (Constructivist Mixed Reality for Design, Education, and Cultural Heritage). This project has two real-world target applications, a virtual heritage reconstruction and an urban planning and architecture project. In the following sections, we describe these applications and discuss the first results of the system in use.

2. Background

We briefly discuss a selection of the most relevant computer graphics/vision approaches for data capture and display. We then present a rapid overview of previous work related to the chosen application domains.

2.1. Model Creation, Realistic Display and VEs

The premise of our modeling approach is the creation of realistic 3D models from images that can be displayed in VEs. In computer graphics, a number of Image-Based Modelling and Rendering techniques have been developed (e.g., [GGSC96, DTM96]). Despite recent advances (e.g., [BBM*01]), these techniques usually require special purpose display methods. Such approaches can be hard to integrate into traditional VR systems, which have numerous software components to handle the complexity of the hardware platform (stereo, tracking, different devices etc.), and are usually created with standard scene-graph APIs such as OpenGL Performer. As a result, the application of such techniques into an integrated VE is rare.

We have chosen to use a modelling-from-images approach (e.g., [DTM96], ImageModeler from REALVIZ (www.realviz.com)). Other forms of 3D scanning (laser scanning etc.) could potentially be used for some of the applications we examine; each approach has different tradeoffs. In the context of situated activity in a VE, we believe that the simplicity and cost of capture from photographs, and the quality of the resulting models justifies our choice. This does not hold true for all applications (for example, where millimeter precision is required). In the long run, we believe that combinations of several different acquisition techniques should be used.

In this work, we have chosen two visual enhancement components which we have integrated into our systems, notably shadows and vegetation. For shadows we have adapted perspective shadow maps [Wil78, SD02]. For the display of vegetation, we use a mixed point-based/polygon rendering technique which allows us to handle complexity efficiently. Point-based rendering [LW85] has recently seen growing interest in the research community (e.g., [RL00, WFP*01]), but has not yet been used in VE's to our knowledge.

The inclusion of 3D spatialised sound [FJT03] is paramount in achieving a truly convincing virtual reality experience [LVK02]. However, integration of 3D sound in virtual reality applications remains limited due to the lack of standardized tools and heavy signal processing costs that do not scale well with the complexity of the auditory scene.

2.2. Education and Design

As VR technology becomes commonplace, there has been a proliferation of VR in fields such as design, education, and entertainment or, in other words, areas where VR applications are more easily available to and accessible by the general public. In the field of education, VEs have been developed to help teach concepts that are hard to learn [RJM*99, BHB99] or difficult to visualise otherwise. In design, VR has been used where conventional media are ill-suited to represent the work processes in ways that make them easy to visualise. In both cases, VR, with its immersive and interactive properties, can offer possibilities and solutions that are otherwise very difficult to obtain.

For these reasons, we have chosen two application domains that relate to learning and working in VR, an archaeological reconstruction and an urban development project. Our choice has been motivated by the fact that both tasks are real-world projects that are currently in progress and in need of high-level tools and presentation means that will speed-up and facilitate the work or help in better dissemination of their cause. In the first case, the Society of Messenian Studies has been involved for years in excavating the temple and the surrounding monuments of ancient Messene and currently reconstructing the archaeological site in order to make it more accessible to visitors (Fig. 2). Additionally, the Foundation of the Hellenic World, a cultural heritage museum which is located in Athens and which uses CAVE-like

† http://www.cs.ucl.ac.uk/create/

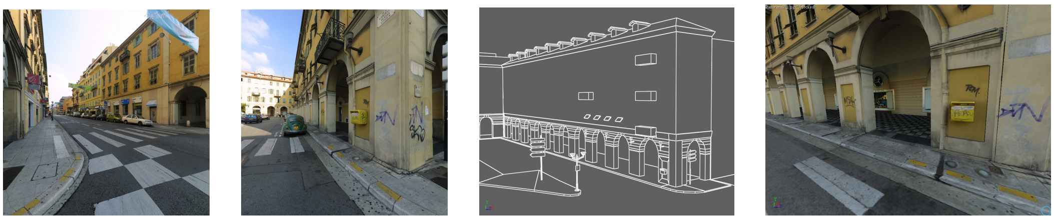

Figure 1: *Two faces of a panorama cubemap used for image modeling. Wireframe of the resulting model and a view with extracted textures.*

and ImmersaDesk^{TM} VR displays to educate children and entertain families, plans to add the final VE in its repertoire of educational and recreational exhibits for visitors.

For the second application, the City of Nice in France and the Urban Community Council (CANCA) have currently commissioned the construction of a tramway that will be integrated in the measures taken to improve urban life by reorganizing and augmenting open space, improving accessibility for pedestrians, stimulating economic life, and helping preserve and display the city's architectural heritage that is endangered by car pollution. The city administration would like to be able to present several designs of the suggested urban intervention, allowing shop owners, local inhabitants, and travellers to evaluate the result from viewpoints they choose (for example, from street level or from the office they work in) or through "guided tours" of typical usage patterns of new infrastructures (for example, how a typical traveller will arrive at a station or cross a street). The architects who have undertaken the development of the project, wish to thoroughly evaluate the effects of the urban intervention on the environment and the landscape of the city, and in particular use interactive VEs to choose amongst different designs.

3. Creating a Realistic Virtual Environment

The creation of textured 3D models from photographs results in low-polygon count models which have a very high level of believability. The resulting 3D environment can be edited using standard 3D modelling tools, and virtual elements can be added as needed.

With such an approach, the textures are view-dependent, since the photograph is valid only from the point it was taken. To counter this problem, we introduce a view-dependent capture and display workflow.

3.1. View-Dependent 3D Models from Images

The data capture workflow starts by choosing a number of viewpoints around which VE activity will take place. Panoramic images are then created by shooting a number of photographs around the viewpoint and "stitching" the images together[‡]. These panoramas are rendered as cubemaps and loaded into an image-modelling program, such as ImageModeler^{TM}. An example of 2 (out of 4) faces of a panorama cubemap are shown in Fig. 1. Other similar systems (e.g., Facade [DTM96], PhotoModeler (www.photomodeler.com)) could also be used. The cameras of images/faces of the cube are calibrated, and 3D models can then be constructed.

The initial phase of the workflow is the construction of an untextured 3D model of the scene, for example Fig. 1 and Fig. 3. This model is the same for all the panoramas used. Textures are then extracted by projecting the image back into object space. Textures are edited at various stages of the process for each view. The initial images of the panorama can be edited in an image-editing program to remove undesirable elements such as people or cars etc. Since we have multiple views, there is an inevitable time and exposure difference in the photographs. To deal with this, manual editing of the panorama images may be needed to equalize colors and to modify the position of shadows. Additional editing (e.g. "clone brushing") is required at the texture extraction phase, for example in regions that were hidden in the original views.

Evidently this process is still not as simple or as rapid as we would like. As an example, the calibration and modelling phase for the urban square (Place Garibaldi in Nice) shown in Fig. 1 and in the accompanying video took about 3 weeks. The area covered, however, is quite large and the amount of geometric detail is significant. Texture extraction, including all the editing phases, takes about a week per panorama.

The result of this process is a single 3D model, and one set of textures per panorama (viewpoint). As discussed later, this allows view-dependent display of the models, resulting in significantly higher quality renderings.

We are actively pursuing ways to improve and simplify the capture process. In particular, we have introduced a novel approach [RD03] which uses projective textures instead of reprojected/resampled textures. The method begins by creating visibility layers for each image, thus facilitating the work

‡ We use a Kaidan head on a tripod and REALVIZ Stitcher^{TM} for this phase, but any standard tool can be used.

of the artist and significantly reducing the texture memory overhead. The fact that texturing takes place entirely in the image-editing program and in the original image space, significantly accelerates the work of the artist.

Nevertheless, problems that arise from the use of view dependent texturing exist and are handled on a case by case basis. We have discovered that a large number of view dependent textures have to be used when the user travels long distances in the VE because in this case the parallax of the images is too large and image coherence is a problem. Depending on the importance of the object being textured a detailed modelling effort can alleviate the problem.

Figure 2: *Upper row: The ancient site of Messene, and the plaster model built by archaeologists. Lower row: photographs taken during the capture process.*

3.2. Data Acquisition for the CREATE Project

The selected sites, located in Greece and in Southern France, were photographed between November 2002 and May 2003. For the cultural heritage application, the site of the temple of Asclepios in the Hellenistic city of Messene (SW Peloponnese in Greece) was chosen due to its cultural, educational, and symbolic significance, since ultimately our goal is to use this VE reconstruction as an educational and presentation tool in a museum. The advantages of the site for capture include the lack of vegetation in the foreground, the existence of numerous easy-to-identify blocks and objects, and a relatively flat terrain (Fig. 2). A disadvantage is that the temple is poorly preserved so the models resulting from the capture process are limited, which means that a large part of the temple must be modelled manually.

On the other hand, a section of the site of Place Massena (Fig. 4) in Nice, France was initially chosen because of its central location in the development of the tramway project. After involving the end-users (the architects and urban planners) more closely in the design through the user task analysis, an additional location was added, that of Place Garibaldi (Fig. 1).

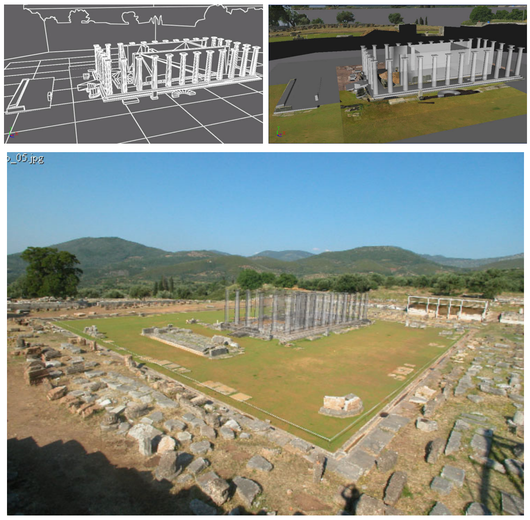

Figure 3: *The Messene site. Upper row, wireframe and shaded model of site and partially reconstructed temple. Lower image, overview.*

3.3. Enhancing Realism of the Virtual Environment

Given the view-dependent textured 3D model, we need a rendering algorithm and an adaptation of display for traditional VE systems. We discuss the two approaches developed for view-dependent texture mapping for immersive systems.

Once the VE has been constructed, we can augment it with virtual elements. For enhanced realism we have added interactive shadow display and vegetation. We have also included spatialized 3D sound which greatly enhances the effect of immersion in the environment. Finally, we also discuss some issues of software engineering and integration that simplified our solution.

View-dependent Texture Mapping

A texture extracted with respect to a certain viewpoint is not valid from others. Nonetheless, the realism provided by textures from photographs is such that the user is often “sufficiently fooled” for non-negligeable motion. In essence enough 3D geometry needs to be reconstructed in the near field, and textures from appropriate viewpoints need to be used.

The principle of view-dependent texture mapping was developed by Debevec et al. [DTM96] as an offline process, who later adapted the approach to interactive viewing [DYB98]. The approach used there was based on projective textures, which require that the geometry be subdivided from each viewpoint used such that hidden parts are separate geometrically.

Due to this restriction and the tool workflow we used, that is ImageModeler^{TM} for modelling and OpenGL^{TM} Performer^{TM} for display, we developed a first approach which is based on texture blending. We have added a node to the Performer scene graph which handles the fact that a single geometry has multiple textures. The draw callback of the node then computes the closest viewpoint and chooses the appropriate texture to apply. In transition regions, blending between textures is applied.

The second approach uses projective textures using layered images [RD03]. As described above, the capture process provides images layered by visibility, thus avoiding the need to subdivide the geometry according to each viewpoint. Each layer of the image is the projective texture corresponding to the appropriate part of the model. A similar blending approach is used for multiple viewpoints. Details of this approach are described in the paper [RD03].

Figure 4: *Left, the VE with the textured version of the above model. Right, the final VE after the addition of virtual elements, realistic shadows and lighting.*

Lighting, Shadows and Vegetation

An important requirement of high-quality rendering is consistent lighting between virtual, inserted elements and the captured geometry corresponding to real objects. For this we have integrated illumination from a sky and sunlight model [PSS99] to achieve consistent lighting of virtual objects. We use efficient acceleration methods to sample the sky hemisphere and achieve the desired effects.

Shadows add a very important dimension to the perception of realism in the environment, particularly when seen on an immersive display. We have developed high-quality shadowing algorithms, which we achieve by using perspective shadow maps [SD02] integrated into a VR system. This integration requires the adaptation of the basic algorithm to the particularities of scene-graph based solutions (see Fig. 4).

To display vegetation we have integrated a point and line-based rendering approach [SD01, DCSD02] for complex vegetation geometry which is in the distance. Again, a new Performer node is created containing the point array. The array exists next to the original geometry in the graph, and when sufficiently small, an appropriate prefix [DCSD02] of the point set is displayed instead of the complex geometry. As a result, we can treat complex models of trees or bushes efficiently permitting interaction with these elements.

Sound

We have designed a 3D sound system that scales well to large numbers of sound sources. Based on spatial and psychoacoustic grouping rules, we dynamically build clusters of sound sources using traditional clustering techniques [LVV03, HS85]. Each cluster is then rendered as a single point source. The sound signal for each cluster is obtained by individually pre-mixing the signals of each source within the cluster. Contrary to prior related work [Her99] our approach performs dynamic clustering that can be used for resource allocation.

We also take into account the content of the source sound signals to enhance the accuracy of our grouping strategy. Our approach currently allows for rendering 5 to 10 times more spatialised sources than current state-of-the-art consumer hardware alone which is limited to 32 or 64 simultaneous 3D sources.

Being able to render a large number of spatialised sound sources gives the opportunity to increase the realism of virtual worlds by including virtual image-sources that account for sound reflection and diffraction or model spatially extended sound sources as collections of point-sources (for additional details on the technique, see [TGD04]). These algorithms are implemented within an audio rendering server that interfaces with our VR library and can either run locally or on a distant machine through a network connection. The core of the server is platform-independent while access to audio hardware is implemented through plug-ins, ensuring portability.

3.4. Software Engineering Issues

The software engineering aspects of this work are considerable. We have based our system on VRCO's CAVElib^{TM} and SGI's OpenGL Performer^{TM} scene graph library. We have used an extensible scripting language XP, originally developed at the Electronic Visualization Laboratory [PIA*98] and extended by the Foundation of the Hellenic World [GCR01]. This language allows the addition of new nodes (such as those for point-based rendering or view-dependent objects), implemented in Performer and directly accessible in the VE scripts.

The benefit of this approach is that these are standard software components, which are provably portable across different VR systems. For instance, most of the examples shown here were initially developed on a PC-based Linux Barco BARON workbench but have been successfully displayed on an SGI IRIX Onyx4-based Reality Center^{TM} and a CAVE-like display without any particular difficulty (see Fig. 5).

The use of such libraries does have its constraints however, since structured scenegraphs typically impose a number of choices concerning rendering modes etc. For example, implementation of the shadow mapping algorithm required

Figure 5: *The VE of the urban planning/architecture application as displayed on an immersive curved-screen display and on a workbench.*

modification of the basic rendering channels, and restricts the efficiency of multiprocessing. Nevertheless, in our experience, the benefit of having working software components for all the VR device issues largely outweighs the difficulties encountered.

4. Adding Interactivity

In addition to the development of visually rich virtual environments, our user-centered approach presupposes the development of usable applications that include a number of interactive capabilities designed to meet real-world user requirements. The design of interactivity for each case has resulted in a set of representative user task scenarios that are currently being evaluated and refined, as described in the next section.

In the first example of the scenario for the final user experience of ancient Messene, the central temple of the site provides a focused activity space that is relatively restricted. The position, scale, and symmetry of the Doric temple concentrates the user's movement in a well-defined area and provides an ideal situation for iterative design, where activity can initially focus on a smaller section and then be replicated multiple times for application onto the whole. According to the user task scenarios developed for the archaeological reconstruction, the novice user will be able to engage in the reconstruction of the temple in a manner that resembles creative child's play with a construction kit; the user may search for the correct pieces that allow to virtually "build", examine, compare, and manipulate parts of the temple as these are found (and accurately captured) in the environment (see Fig. 6), and, during this process, actively explore various levels of reconstruction for different historical periods. Similarly, for an archaeology scholar who intends to restore the ancient monument, the goal is to offer the possibility to try varied reconstruction hypotheses and choose the most plausible one. Such an activity usually involves the assemblage of several alternative but partial hypotheses for each structural part of the monument: ground plan, elevation, number and position of columns, size and form of main building, height and material of the roofing, etc. Hence, the kinds of actions a domain expert would choose to do include the manipulation of all available architectural members (resizing, scaling, positioning), the repair/restoration of damaged members with the use of primary material (virtual plaster, wood, stone), the readaptation of general data such as the column distance, the monument height, the thickness of the walls, and so on.

Similar activity has been designed for the urban development case where the addition of interactive features may allow users to dynamically investigate alternatives by directly manipulating environmental elements (pedestrian crossings, stations, buildings, vegetation, etc.). The architects of the tramway project in Nice will use the VE's interactive features to make judgments, examine alternative development scenarios, and experiment with different possibilities and solutions during their virtual prototyping process (observe, for instance, the aesthetic and functional effect of urban interventions). These processes, if judged successful for part of a site, may be repeated for a larger area. It is believed that by such direct manipulation in an environment, which preserves close-to-real spatial relationships, users can better understand the effect of specific actions, achieving better overall comprehension of the task and, through the active iterative approach, gain a better sense of the relative dimensions or interrelationships between the various elements under examination.

Overall Process

As mentioned above, for the Garibaldi scene (which is an entire square of a city), the overall photography, calibration, modelling and texturing required about 2 person-months. The addition of interactivity is performed via the creation of XP scripts, and required 1-2 weeks. A trained artist can perform all of these steps without the need for a programmer. Nonetheless, our system is still a prototype, so in practice many of the steps were quite cumbersome. Many aspects

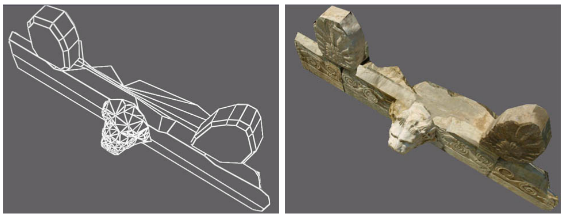

Figure 6: *A view of the 3D model of the* sima *(a part of the temple) created from a number of different images. This is one of the objects that will be manipulated by the users during their interaction with the virtual environment.*

could be automated and the user interface could be much simpler. Given our accumulated experience, we believe that the time required, with the current system, could be reduced by half.

5. User-Centered Design and Evaluation

The practice of applying user-centered design and evaluation methods in virtual reality is currently gaining ground as illustrated by the human factors research in the field [SMK98] and the recent application examples in training and education [MNW03, NCW02]. Involving users from the early stages of this project is the central tenet of our approach, which entails designing *for* and designing *with* the users of the resulting virtual environments [SaFAD97].

Our methodology is largely based on Gabbard, Hix and Swan's work [GHS99], who propose a structured framework for the design and evaluation of user activity in VEs that includes user task analysis, heuristic and usability guidelines, formative, and summative evaluation. We have extended this framework to take into account the specific application domains of education and design, and the different types of users that our work targets, including children (a challenging case that requires particular attention and a more qualitative approach).

We have augmented this approach by taking audiovisual fidelity into account in the design and evaluation of our VE's. In the user requirements analysis we try to identify cases in which increased realism is important. An example is the design of outdoor spaces by architects. In this case the perceived realism of the VE simulation is paramount in their ability to apprehend the overall feeling and ambiance of the environment being designed. Accurate and full sound simulation is an integral part of this approach. In addition, the ability to interactively manipulate very realistic elements (vegetation, view-dependent textured objects etc.) helps them make more informed decisions. For future experiments, the evaluation of the effect of realism will be central.

This framework ensures that users do not only assume the role of testers after the development of the environments is completed, but are involved early in the design process in order to guarantee that the developed VEs meet their needs. This is best illustrated through our work with both the archaeologists and educators in the case of the archaeological reconstruction and the architects and urban planners in the case of the urban tramway project.

5.1. User Task Analysis

For the purpose of designing the case studies for this project, a user task analysis was performed which includes detailed user task descriptions in the form of scenarios for each case and each user group. Prior to the needs analysis, the target user groups for each case were identified and defined. For both domains, experts (restoration architects, archaeologists, archaeology students, as well as architects and urban planners) and novices (children, tourists, museum visitors, as well as city administration and the general public) were selected and interviewed. As our interviews indicated, prospective users of either application generally felt that, when carrying out the tasks we described, they are usually confronted with a multitude of options that cannot be easily examined, interpreted or placed into context. Additionally, the dissociation of information between a real-world problem and conventional computer tools makes it difficult to visualise and situate the gained knowledge of a situation, leading to a variety of misconceptions and false assumptions. Even for experts, a building activity is always a challenging task since a number of different interpretations and hypotheses exist that have to be evaluated, while the tools available to specialists today are usually low-tech, low-accuracy passive environments that cannot give the correct impression of the context.

The results of this needs analysis process have informed the design and development of our prototypes, as well as the design of the formative evaluation studies that are currently underway.

5.2. Preliminary Formative Studies

Formative evaluation studies have been designed in phases. The first in a sequence of testing phases, which has been completed, involved the development of exploratory studies that were conducted, for the education case, with young students between the age of 7 and 12, and for the urban planning case, with the architects of the project. The main objective of these early experiments was to explore the medium and the method, namely to clarify technical issues concerning the methodology for testing, to develop the method of working with the different users for each case study, and to practice the analytical framework.

The tasks included simple building activities that involved the selection, placement, and manipulation of virtual objects (see Fig. 7). When reviewing the case studies, we looked for critical incidents or, *contradictions*, which occurred while the user interacted with the system, both on a technical and a conceptual level. The analysis of these incidents was guided by the framework of Activity Theory [Nar96] [BHB99].

Figure 7: *A young student interacting with an early version of one of the CREATE environments during the exploratory user studies in a cubic immersive display.*

Finally, in what concerns the technical and usability evaluation of the prototype environment, individual experiments were performed in order to assess the technical choices made in the project and the effectiveness of the developed algorithms. Additionally, issues that hindered user interaction within the prototype environment were recorded and are being addressed in the development of further prototypes.

5.3. Conclusions and Further Work

Virtual environments, despite technological advancement, suffer either from a lack of realism or a low degree of interactivity. We believe that both these properties are important and must be combined in VR applications that involve archaeological reconstruction and education or urban planning, where expert and novice end-user needs must be addressed. The user-centered approach described in this paper, along with the techniques for capturing and enhancing realism and adding interactivity, attempt to make progress in this direction.

The introduction of high-fidelity visual and auditory environments is a significant step in allowing VE's to be used in real-world tasks. The addition of view-dependent texturing for realistic and interactive viewing of existing real-world objects in VE's, realistic vegetation, shadows and 3D sound, all contribute to an increased sensation of realism and immersion. Coupled with interactivity, we believe that these elements will permit users to make more informed decisions and improve their capacity to learn and design.

The exploratory pilot studies carried out thus far have been very useful in providing directions for the design of the next studies, while user involvement and response has been essential in this process. In terms of usability, the studies indicated that the complex user activity in the VEs requires seamless multi-modal interaction methods that will facilitate the users in better performing the tasks. We have decided to accommodate this need by extending the interactive experience to the other senses -such as touch- by developing haptic interfaces that will allow the designer/learner/visitor/scholar to dynamically build the virtual environment in a natural manner. In terms of enhancing realism, the addition of virtual people is another aspect that is considered important, especially for the urban planning case, and is currently underway.

Finally, the user studies have only just begun and will continue, both on a formative and on a summative level, so as to ensure that the resulting application environments meet their users' needs and provide virtual reality tools that address real-world situations.

6. Acknowledgments

CREATE is a 3-year RTD project funded by the 5th Framework Information Society Technologies (IST) Programme of the European Union (IST-2001-34231), http://www.cs.ucl.ac.uk/create/. The authors wish to thank all the individuals from the partner institutions (University College London, INRIA, Foundation of the Hellenic World, REALVIZ SA., University of Cyprus, PERCRO - Scuola Superiore S.Anna, and CSTB - Centre Scientifique et Technique du Batiment) that have contributed to the development of the proposal and the project. In particular, thanks are due to Alexandre Olivier-Mangon for model creation, texturing and for shooting and editing the accompanying video, to Saki Gaitatzes for proof-reading and his comments, and to Celine Loscos for her project coordination efforts. Thanks also to Alias|Wavefront for a generous donation of its Maya software.

References

[BBM*01] BUEHLER C., BOSSE M., MCMILLAN L., GORTLER S. J., COHEN M. F.: Unstructured lumigraph rendering. In *SIGGRAPH 2001, Computer Graphics Proc.* (2001), Annual Conference Series, pp. 425–432. 1, 2

[BHB99] BARAB S. A., HAY K., BARNETT M. G.: Virtual solar system project: Building understanding through model building. In *Annual Meeting of the American Educational Research Association* (Montreal, Canada, April 1999), AERA. 2, 8

[DCSD02] DEUSSEN O., COLDITZ C., STAMMINGER M., DRETTAKIS G.: Interactive visualization of complex plant ecosystems. In *Proc. of IEEE Visualization 02* (October 2002), IEEE. 5

[DTM96] DEBEVEC P., TAYLOR C., MALIK J.: Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. In *Proc. SIGGRAPH 96* (August 1996), pp. 11–20. 1, 2, 3, 4

[DYB98] DEBEVEC P., YU Y., BORSHUKOV G.: Efficient view-dependent image-based rendering with projective texture-mapping. In *Rendering Techniques '98* (9th EG workshop on Rendering, June 1998). 4

[FJT03] FUNKHOUSER T., JOT J.-M., TSINGOS N.: Survey of methods for modeling sound propagation in interactive virtual environment systems. *accepted for publication in Presence.* (2003). 2

[GCR01] GAITATZES A., CHRISTOPOULOS D., ROUSSOU M.: Reviving the past: cultural heritage meets virtual reality. In *Proceedings of the 2001 conference on Virtual reality, Archeology, and Cultural Heritage* (2001), ACM Press, pp. 103–110. 5

[GGSC96] GORTLER S. J., GRZESZCZUK R., SZELISKI R., COHEN M. F.: The lumigraph. In *SIGGRAPH 96 Conference Proc.* (Aug. 1996), Annual Conference Series, pp. 43–54. 1, 2

[GHS99] GABBARD J. L., HIX D., SWAN II J. E.: User-centered design and evaluation of virtual environments. *IEEE Computer Graphics and Applications 19* (November/December 1999), 51–59. 7

[Her99] HERDER J.: Optimization of sound spatialization resource management through clustering. *The Journal of Three Dimensional Images, 3D-Forum Society 13*, 3 (Sept. 1999), 59–65. 5

[HS85] HOCHBAUM D. S., SCHMOYS D. B.: A best possible heuristic for the *k*-center problem. *Mathematics of Operations Research 10*, 2 (May 1985), 180–184. 5

[LVK02] LARSSON P., VÄSTFJÄLL D., KLEINER M.: Better presence and performance in virtual environments by improved binaural sound rendering. *proceedings of the AES 22nd Intl. Conf. on virtual, synthetic and entertainment audio, Espoo, Finland* (June 2002), 31–38. 2

[LVV03] LIKAS A., VLASSIS N., VERBEEK J.: The global k-means clustering algorithm. *Pattern Recognition 36*, 2 (2003), 451–461. 5

[LW85] LEVOY M., WHITTED T.: The use of points as a display primitive. In *CS Department, University of North Carolina at Chapel Hill* (Jan. 1985), TR 85-022. http://www-graphics.stanford.edu/papers/points/. 2

[MNW03] MARSHALL E., NICHOLS S., WILSON J. R.: Interactivity, Control of Movement and Realism: Establishing the Factors Influencing Virtual Reality Training. In *Proceedings of HCI International 2003* (June 2003), vol. II, Lawrence Erlbaum Associates, Inc, pp. 1188–1192. 7

[Nar96] NARDI B. A.: *Context and Consciousness: Activity Theory and Human-Computer Interaction*. MIT Press, Cambridge, Massachusetts, 1996. 8

[NCW02] NEALE H., COBB S., WILSON J. R.: A Front-End Approach to the User-Centred Design of VEs. In *Proceedings of IEEE Virtual Reality 2002* (2002), IEEE Computer Society. 7

[PIA*98] PAPE D., IMAI T., ANSTEY J., ROUSSOU M., DEFANTI T.: XP: An authoring system for immersive art exhibitions. In *4th International Conference on Virtual Systems and Multimedia (VSMM)* (1998), Gifu, Japan. 5

[PSS99] PREETHAM A. J., SHIRLEY P., SMITS B.: A practical analytic model for daylight. In *SIGGRAPH* (1999). 5

[RD03] RECHE A., DRETTAKIS G.: View dependent layered projective texture maps. In *Proceedings of Pacific Graphics 2003* (October 2003), Rokne J., Klein R.,, Wang W., (Eds.), IEEE Press. 3, 5

[RJM*99] ROUSSOS M., JOHNSON A. E., MOHER T. G., LEIGH J., VASILAKIS C. A., BARNES C. R.: Learning and building together in an immersive virtual world. *PRESENCE journal 8*, 3 (June 1999), 247–263. 2

[RL00] RUSINKIEWICZ S., LEVOY M.: QSplat: A multiresolution point rendering system for large meshes. In *Proc. ACM SIGGRAPH* (2000), pp. 343–352. 2

[SaFAD97] SCAIFE M., AN F. ALDRICH Y. R., DAVIES M.: Designing for or designing with? Informant design for interactive learning environments. In *Proceedings of CHI '97: Human Factors in Computing Systems* (1997), ACM, New York. 7

[SD01] STAMMINGER M., DRETTAKIS G.: Interactive sampling and rendering for complex and procedural geometry. In *Rendering Techniques 2001* (2001), 12th EG workshop on Rendering. 5

[SD02] STAMMINGER M., DRETTAKIS G.: Perspective shadow maps. In *Proceedings of ACM SIGGRAPH 2002* (July 2002), Hughes J., (Ed.), Annual Conference Series, ACM Press/ ACM SIGGRAPH. 2, 5

[SMK98] STANNEY K. M., MOURANT R. R., KENNEDY R. S.: Human factors issues in virtual environments: A review of the literature. *PRESENCE journal 7*, 4 (1998), 327–351. 7

[TGD04] TSINGOS N., GALLO E., DRETTAKIS G.: Perceptual audio rendering of complex virtual environments. In *Proc. ACM SIGGRAPH 2004 (Special Issue of ACM Transactions on Graphics)* (2004). 5

[WFP*01] WAND M., FISCHER M., PETER I., AUF DER HEIDE F. M., STARSSER W.: The randomized z-buffer algorithm: Interactive rendering of highly complex scenes. In *Proc. ACM SIGGRAPH* (2001). 2

[Wil78] WILLIAMS L.: Casting curved shadows on curved surfaces. *Computer Graphics (Proc. of SIGGRAPH 78) 12*, 3 (1978), 270–274. 2