INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team REVES

Rendering and Virtual Environments with Sound

Sophia Antipolis - Méditerranée

THEME COG

Activity Report

2007

Table of contents

1. Team **1**
2. Overall Objectives **1**
2.1. General Presentation 1
2.2. Highlights of the year 2
3. Scientific Foundations **2**
3.1. Rendering 2
3.1.1. Plausible Rendering 2
3.1.1.1. Alternative representations for complex geometry 2
3.1.1.2. Plausible audio rendering 3
3.1.2. High Quality Rendering Using Simulation 3
3.1.2.1. Non-diffuse lighting 3
3.1.2.2. Visibility and Shadows 4
3.1.2.3. Radiosity 4
3.1.2.4. High-quality audio rendering 4
3.2. Virtual and Augmented Environments with Sound 5
3.2.1. Efficient and Simple Relighting 5
3.2.2. Enriching virtual environments with sound 6
3.2.3. Interaction and Visual Paradigms for Virtual and Augmented Environments 6
4. Software **7**
4.1. AURELI: AUdio REndering LIbrary/AUDILE 7
4.2. OgreVR, OgreAL and OgreVRScript 7
4.3. LibSL - Simple Library for Graphics 8
5. New Results **8**
5.1. Plausible Image Rendering 8
5.1.1. Antiradiance 8
5.1.2. Tile-Trees 9
5.1.3. Compact Texture Synthesis 9
5.1.4. Texturing from Photographs 10
5.1.5. Mesh puzzle 10
5.1.6. Filtered Tile Maps 11
5.1.7. Compressed Random Accessed Trees for Spatially Coherent Data 11
5.1.8. Interactive image editing 11
5.1.9. Surface Solid Texture Synthesis from Slices 12
5.1.10. An Interactive Perceptual Rendering Pipeline using Contrast and Spatial Masking 12
5.2. Plausible Audio Rendering 13
5.2.1. 3D Audio matting and re-rendering 13
5.2.2. Instant sound scattering 13
5.2.3. Reverberation graphs 15
5.2.4. Processing Massive Impact Sounds 16
5.3. Virtual Environments with Sound 16
5.3.1. Progressive Perceptual Audio Rendering of Complex Scenes 16
5.3.2. Contact sound shading 17
5.3.3. Auditory Masking for Scalable VOIP Bridges 17
5.3.4. Design and Evaluation of a Real-World Virtual Environment for Architecture and Urban Planning 17
5.3.5. Role of semantic vs spatial congruency in a bimodal go/no-go task 18
6. Contracts and Grants with Industry **19**
6.1. Alias|Wavefront 19
6.2. Eden Games 19

6.3. CIFRE CSTB 19
6.4. Renault 19
7. Other Grants and Activities 19
7.1. Regional/Local Projects 19
7.1.1. Collaboration with CNRS and IRCAM 19
7.1.2. Collaboration with CSTB Sophia-Antipolis 19
7.1.3. The workbench platform 20
7.2. European Projects 20
7.3. Visiting Researchers 21
7.4. Bilateral Collaborations 21
7.4.1. France-China 21
7.4.2. France-Germany 21
7.4.3. France-United States of America 21
7.4.4. France-Italy 21
8. Dissemination 21
8.1. Participation in the Community 21
8.1.1. Program Committees 21
8.1.2. Distinctions 22
8.1.3. Invited Talks 22
8.1.4. Thesis Committees 22
8.1.5. COST and CUMIR 22
8.1.6. Web server 22
8.2. Teaching 22
8.2.1. University teaching 22
8.2.2. PhD Thesis Completed and Continuing 23
8.3. Participation at conferences 23
8.3.1. Presentations at Conferences 23
8.3.2. Participation at Conferences and Workshops 23
8.4. Demonstrations and Press 23
8.4.1. Demonstrations 23
8.4.2. Press 23
9. Bibliography 23

1. Team

Head of project-team
George Drettakis [Research Director (DR) Inria, HdR]

Vice-head of project team
Nicolas Tsingos [Research Associate (CR) Inria]

Member of project team
Sylvain Lefebvre [Research Associate (CR) Inria]

Postdoctoral Fellows
Carsten Dachsbacher [Marie-Curie Post-doctoral fellowship]
Efstathios Stavrakis [INRIA Post-doctoral fellowship]

Technical staff
David Geldreich [Part time IR, Workbench/DREAM]
David Grelaud [Associated Engineer]
Fanourios Moiratis [Expert Engineer/Modeller]

Ph.D. students
Emmanuel Gallo [CIFRE/CSTB]
Nicolas Bonneel [CROSSMOD project]
Cécile Picard [INRIA/EdenGames collaboration]
Marcio Cabral [INRIA CORDIS scholarship]

Research scientist (external collaborators)
Frédo Durand [MIT CSAIL]
Hugues Hoppe [Microsoft Research]
Maria Roussou [University College London]
Marc Stamminger [Prof. University of Erlangen]
Isabelle Viaud-Delmon [CNRS]
Clara Suied [CNRS]

Administrative assistants
Aicha Venturini [Secretary (SAR) Inria, until May]
Sophie Honnorat [Secretary (SAR) Inria, since September]
Monique Varlet [CROSSMOD assistant]

Interns
Thomas Moeck [Master Erlangen University]
Christian Eisenacher [Master Erlangen University]
Nader Salman [Master University of Nice]

2. Overall Objectives

2.1. General Presentation

Images, often accompanied by sound effects, have become increasingly present in our everyday lives; this has resulted in greater needs for content creation. Despite the fact that many traditional means exist, such as photography, artistic graphic design, audio mixing, they typically still remain the reserve of the expert, and require significant investment in time and expertise.

Our main interest is computer image and sound synthesis, with an emphasis on automated methods. Our main goals include the simplification of the tasks required for the production of sound and images, as well as the development of new techniques for their generation.

The application domain is vast. It ranges from audiovisual production, which typically requires long, offline computation to obtain high quality results, all the way to real-time applications such as computer games or virtual reality, for which the main consideration is to guarantee 60 frames per second frame rates, or, in general the reduction of latency to user reaction.

The process of generation of images and sound, generally called *rendering* is our primary interest; our second main interest are virtual environments (VE's) as well as augmented (AE's) or mixed environments (ME's), that is scenes containing both real objects (often digitized) as well as purely synthetic objects. We are interested in both the generation and the interaction with these environments. We use the term virtual environments for scenes with a certain degree of interactivity, potentially in a semi-immersive (stereo and tracking, workbench) or immersive (CAVE, RealityCenter) context.

2.2. Highlights of the year

In 2007, George Drettakis received the Outstanding Technical Contributions Award from the Eurographics Association. He also was elected to a Fellowship of the Association in recognition of his contribution to the field.

3. Scientific Foundations

3.1. Rendering

Keywords: *high-quality rendering*, *image rendering*, *plausible rendering*, *rendering*, *sound rendering*.

We consider plausible rendering to be a first promising research direction, both for images and for sound. Recent developments, such as point rendering, image-based modeling and rendering, and work on the simulation of aging indicate high potential for the development of techniques which render *plausible* rather than extremely accurate images. In particular, such approaches can result in more efficient renderings of very complex scenes (such as outoors environments). This is true both for visual (image) and sound rendering. In the case of images, such techniques are naturally related to image- or point-based methods. It is important to note that these models are becoming more and more important in the context of network or heterogeneous rendering, where the traditional polygon-based approach is rapidly reaching its limits.

Another research direction of interest is realistic rendering using simulation methods, both for images and sound. In some cases, research in these domains has reached a certain level of maturity, for example in the case of lighting and global illumination. For some of these domains, we investigate the possibility of technology transfer with appropriate partners. Nonetheless, certain aspects of these research domains, such as visibility or high-quality sound still have numerous and interesting remaining research challenges.

3.1.1. Plausible Rendering

3.1.1.1. Alternative representations for complex geometry

The key elements required to obtain visually rich simulations, are sufficient geometric detail, textures and lighting effects. A variety of algorithms exist to achieve these goals, for example displacement mapping, that is the displacement of a surface by a function or a series of functions, which are often generating stochastically. With such methods, it is possible to generate convincing representations of terrains or mountains, or of non-smooth objects such as rocks. Traditional approaches used to represent such objects require a very large number of polygons, resulting in slow rendering rates. Much more efficient rendering can be achieved by using point or image based rendering, where the number of elements used for display is view- or image resolution-dependent, resulting in a significant decrease in geometric complexity.

Such approaches have very high potential. For example, if all object can be rendered by points, it could be possible to achieve much higher quality local illumination or shading, using more sophisticated and expensive algorithms, since geometric complexity will be reduced. Such novel techniques could lead to a complete replacement of polygon-based rendering for complex scenes. A number of significant technical challenges remain to achieve such a goal, including sampling techniques which adapt well to shading and shadowing algorithms, the development of algorithms and data structures which are both fast and compact, and which can allow interactive or real-time rendering. The type of rendering platforms used, varying from the high-performance graphics workstation all the way to the PDA or mobile phone, is an additional consideration in the development of these structures and algorithms.

Such approaches are clearly a suitable choice for network rendering, for games or the modelling of certain natural object or phenomena (such as vegetation, e.g. Figure 1, or clouds). Other representations merit further research, such as image or video based rendering algorithms, or structures/algorithms such as the "render cache" [40], which we have developed in the past, or even volumetric methods. We will take into account considerations related to heterogeneous rendering platforms, network rendering, and the appropriate choices depending on bandwith or application.

Point- or image-based representations can also lead to novel solutions for capturing and representing real objects. By combining real images, sampling techniques and borrowing techniques from other domains (e.g., computer vision, volumetric imaging, tomography etc.) we hope to develop representations of complex natural objects which will allow rapid rendering. Such approaches are closely related to texture synthesis and image-based modeling. We believe that such methods will not replace 3D (laser or range-finger) scans, but could be complementary, and represent a simpler and lower cost alternative for certain applications (architecture, acheology etc.).

We are also investigating methods for adding "natural appearance" to synthetic objects. Such approaches include *weathering* or *aging* techniques, based on physical simulations [28], but also simpler methods such as accessibility maps [37]. The approaches we intend to investigate will attempt to both combine and simplify existing techniques, or develop novel approaches which are based on generative models based on observation of the real world.

3.1.1.2. Plausible audio rendering

Similar to image rendering, plausible approaches can be designed for audio rendering. For instance, the complexity of rendering high order reflections of sound waves makes current geometrical approaches inappropriate. However, such high order reflections drive our auditory perception of "reverberation" in a virtual environment and are thus a key aspect of a plausible audio rendering approach. In complex environments, such as cities, with a high geometrical complexity, hundreds or thousands of pedestrians and vehicles, the acoustic field is extremely rich. Here again, current geometrical approaches cannot be used due to the overwhelming number of sound sources to process. We study approaches for statistical modeling of sound scenes to efficiently deal with such complex environments. We also study perceptual approaches to audio rendering which can result in high efficiency rendering algorithms while preserving visual-auditory consistency if required.

3.1.2. High Quality Rendering Using Simulation

3.1.2.1. Non-diffuse lighting

A large body of global illumination research has concentrated on finite element methods for the simulation of the diffuse component and stochastic methods for the non-diffuse component. Mesh-based finite element approaches have a number of limitations, in terms of finding appropriate meshing strategies and form-factor calculations. Error analysis methodologies for finite element and stochastic methods have been very different in the past, and a unified approach would clearly be interesting. Efficient rendering, which is a major advantage of finite element approaches, remains an overall goal for all general global illumination research.

For certain cases, stochastic methods can be efficient for all types of light transfers, in particular if we require a view-dependent solution. We are also interested both in *pure* stochastic methods, which do not use finite element techniques. Interesting future directions include filtering for improvement of final image quality as well as beam tracing type approaches [38] which have been recently developed for sound research.

Figure 1. Plausible rendering of an outdoors scene containing points, lines and polygons [27], representing a scene with trees, grass and flowers. We can achieve 7-8 frames per second compared to tens of seconds per image using standard polygonal rendering.

3.1.2.2. Visibility and Shadows

Visibility calculations are central to all global illumination simulations, as well as for all rendering algorithms of images and sound. We have investigated various global visibility structures, and developed robust solutions for scenes typically used in computer graphics. Such analytical data structures [32] [31], [30] typically have robustness or memory consumption problems which make them difficult to apply to scenes of realistic size.

Our solutions to date are based on general and flexible formalisms which describe all visibility event in terms of generators (vertices and edges); this approach has been published in the past [29]. Lazy evaluation, as well as hierarchical solutions are clearly interesting avenues of research, although are probably quite application dependent.

3.1.2.3. Radiosity

For purely diffuse scenes, the radiosity algorithm remains one of the most well-adapted solutions. This area has reached a certain level of maturity, and many of the remaining problems are more technology-transfer oriented We are interested in interactive or real-time renderings of global illumination simulations for very complex scenes, the "cleanup" of input data, the use of application-dependent semantic information and mixed representations and their management.

Hierarchical radiosity can also be applied to sound, and the ideas used in clustering methods for lighting can be applied to sound.

3.1.2.4. High-quality audio rendering

Our research on high quality audio rendering is focused on developing efficient algorithms for simulations of geometrical acoustics. It is necessary to develop techniques that can deal with complex scenes, introducing efficient algorithms and data structures (for instance, beam-trees [33] [38]), especially to model early reflections or diffractions from the objects in the environment.

Validation of the algorithms is also a key aspect that is necessary in order to determine important acoustical phenomena, mandatory in order to obtain a high-quality result. Recent work by Nicolas Tsingos at Bell Labs [34] has shown that geometrical approaches can lead to high quality modeling of sound reflection and diffraction in a virtual environment (Figure 2). We will pursue this research further, for instance by dealing with more complex geometry (e.g., concert hall, entire building floors).

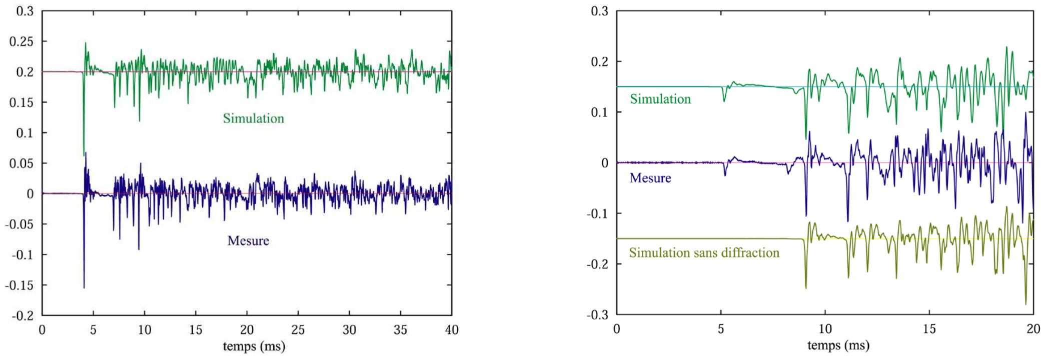

Figure 2. A comparison between a measurement (left) of the sound pressure in a given location of the "Bell Labs Box", a simple test environment built at Bell Laboratories, and a high-quality simulation based on a beam-tracing engine (right). Simulations include effects of reflections off the walls and diffraction off a panel introduced in the room.

Finally, several signal processing issues remain in order to properly and efficiently restitute a 3D soundfield to the ears of the listener over a variety of systems (headphones, speakers). We would like to develop an open and general-purpose API for audio rendering applications. We already completed a preliminary version of a software library: AURELI [39].

3.2. Virtual and Augmented Environments with Sound

Keywords: *augmented environments*, *augmented reality*, *auralisation*, *inverse rendering*, *re-lighting*, *sound "ambiance"*, *virtual environments*, *virtual reality*.

The second major research direction of our group is on virtual, augmented or mixed environments, which include both visual and sound representations. We are mainly interested in interactive environments, permitting the user to create and maniuplate scenes consisting of both real and synthetic objects. As a first step, we consider *real* objects to be digitised representations of reality, rather than the real world.

Our first goal is to apply and adapt our rendering expertise, presented in the previous paragraphs to virtual and augmented reality. There are three areas in which we concentrate our efforts: consistent lighting between real and synthetic illumination, for shadows and reflections, enriching virtual and augmented environments with sound, in a consistent manner and finally appropriate interaction and visual paradigms for virtual and augmented environments.

3.2.1. Efficient and Simple Relighting

We wish to develop relighting and consistent real/virtual lighting methods which have simple input requirements: i.e., a small number of input images, and the smallest number of restrictions on the lighting conditions. The goal is to get high quality results for both interior and outdoors environments. To achieve these goals, we investigate ways to extract approximate reflectances in real scenes, potentially using scene or image statistics, and by including some level of user interaction in the process. For efficient display, texture capacities of modern graphics hardware will definitely be advantageous.

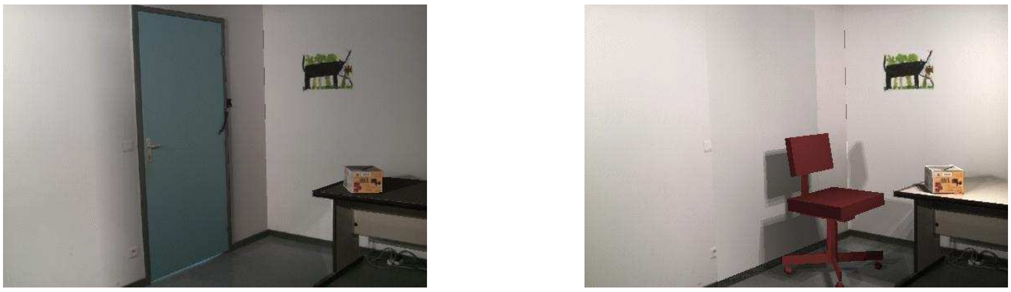

Figure 3. (a) Original conditions (b) The door has been removed virtually, and a virtual object and light have been added (method of [36])

Our previous work on interior relighting (Figure 3) has given satisfactory solutions, allowing us to add virtual object with consistent lighting, but implies severe restrictions on the lighting conditions of input images [35], [36]. Such approaches are based on the creation of "shadow free" base textures using heuristics, and a relatively precise reconstruction of the geometry. For outdoors scenes, geometric complexity and the fact that lighting conditions cannot be easily manipulated render such approaches less appropriate. However, some of the techniques developed can be applied, and we believe that the key is to combine automated techniques with user interaction at the various stages of the process.

The long-term goal is to turn on a video camera in a scene (potentially with partially pre-reconstructed geometry), and be able to add virtual objects or light sources interactively in a consistent manner into the video stream. Relighting could also be achieved in this manner, or using semi-transparent glasses or headsets. Applications of such an approach are numerous, for archeology, architecture and urban planning, special effects, manufacturing, design, training, computer games etc.

This long term vision will require a way to smoothly vary from low-quality methods [35], [36] to high quality approaches [41], in a manner which is much less complex in terms of capture, processing for relighting and (re)rendering.

3.2.2. Enriching virtual environments with sound

Consistent rendering of real and synthetic sounds is a key aspect for virtual reality applications. Solving the problem would make it possible to mix natural sounds with synthesized spatial audio for augmented reality applications. This can be used to enrich the natural soundscape with additional auditory information through wearable devices (e.g., virtual museums, etc.). Another application would be to provide auditory feedback to visually-impaired people while preserving their natural auditory perception.

Another future direction of research is active control of rooms and listening spaces. Such control can be achieved by coupling microphones and speaker arrays and allow for modifying the natural acoustical properties of the space (e.g., reverberation time) in real-time. Such technologies have already been used to improve acoustics in concert halls that, for a variety of reasons, do not sound as good as designed for. They appear to be promising for VR/AR applications. However, existing techniques yet have to be improved to be applied in this context.

3.2.3. Interaction and Visual Paradigms for Virtual and Augmented Environments

The use of immersive or semi-immersive systems opens a large number of new types of interaction with virtual or augmented environments. There is a vast body of research on interfaces for 3D environments, and in particular for immersive systems. Our focus will be on specific interfaces, interaction or visual paradigm problems which inevitably appear in the course of our research. When necessary, we will work with complementary partners in Computer-Human Interaction to find solutions to these problems.

One question we consider important is finding appropriate interface paradigms which replace 2D (menu or button-based) interfaces both in the context of the actual rendering research process and for the applications we investigate. Despite significant previous work in the domain, there is yet to be a standard which has been widely adopted. It may be that the lack of standard interfaces is part of the reason why immersive systems are not being adopted as widely nor as rapidly as their inventors would have hoped.

In terms of visual representation, non-photorealistic (NPR) or expressive, renderings are an interesting avenue of investigation. In particular, NPR can allow abstraction of unimportant details and more efficient communication of certain concepts. Since a number of the algorithms developed are based on inherently 2D drawing, their transposition to immersive, stereo-display environments poses a number of very interesting and challenging questions. There are also some applications domains, for example archeology or architecture, where drawing-style renderings are part of the current workflow, and which will naturally fit into a EVs adapted to these domains. Virtual storytelling is another domain in which NPR has a natural application.

Immersive, stereo-based systems seem a well-adapted platform for more intuitive interactive modelling in 3D. The development of efficient and flexible structures such as procedural point-based representations, or rapid aging techniques in a true 3D context could result in systems which are much more efficient that 2D displays, in which the sensation of 3D depth and immersion is missing.

Finally, the inclusion of spatialised sound for 3D interfaces is clearly a promising research direction. The benefit of consistent 3D sound is evident, since it results in better spatial perception for the user, can help for example in determining spatial or visibility relationships, resulting in improved usability. The actual inclusion of sound effects or sound metaphors in interface design is clearly an interesting challenge.

4. Software

4.1. AURELI: AUdio REndering LIbrary/AUDILE

Participants: Nicolas Tsingos, Emmanuel Gallo, Thomas Moeck, Nicolas Bonneel.

REVES is developing an API, AURELI (AUdio REndering LIbrary), as a tool supporting our research in acoustic modeling and audio rendering. Several prototype algorithms for sound spatialization, geometrical and statistical reverberation modeling, sound source clustering and audio rendering server have been implemented using AURELI's core functionnalities or as an extension to the API itself. Core fonctionalities include audio i/o plug-ins, audio buffer handling and basic signal processing. Higher level functions perform geometrical processing and audio rendering on a variety of restitution systems. AURELI is a cross-platform, object oriented, C++ API. It runs on LINUX/Windows/IRIX and also features primitives for parallel signal processing on multi-processor systems and network communication (used for instance to design audio rendering servers).

We are investigating possibilities for public release of the API as a tool for researchers in acoustics and virtual/augmented reality.

AUDILE is the codename for the clustering, perceptual masking and scalable processing technology, part of which has been transferred to EdenGames. As part of the work on the European project CROSSMOD, an independent library has been created which interfaces to both OpenAL and AURELI in a transparent manner. This software is available on the CROSSMOD repository server.

4.2. OgreVR, OgreAL and OgreVRScript

Participants: David Geldreich, Nicolas Bonneel, George Drettakis.

In the context of his DREAM "mission" David Geldreich has written a framework (OgreVR) to use Ogre3D, an open-source graphics rendering engine, on virtual reality platform. OgreVR abstracts displays/devices setups through a configuration file. OgreVR/Ogre3D is the new software basis for our VR platform and is currently being used by several projects within the group. It is avalaible on the REVES project CVS and is available in SDK form in the CROSSMOD SVN.

As part of his internship, J. Gueytat, developed *OgreAL*, the integration of OpenAL into Ogre3D. This involves two main aspects: A binding for sound sources and their control, which is accessible to the graphics application, and an update phase when graphics objects are moved so that their sound equivalent is also moved. This integration lets us test and develop sound and graphics algorithms within the OgreVR framework.

Concrete experiments require not only 3D models, but also a specific interaction behavior. The goal is to allow also non-CG-experts (e.g., Neuroscience partners of the CROSSMOD project) to specify the interaction in a 3D environment without the need to actually program and compile the code.

During his summer internship, R. Armati, in collaboration with D. Geldreich, developed a set of specifications and requirements for such an authoring tool, in the context of the CROSSMOD project. Different alternatives have been then evaluated. The final decision was to take an existing scripting language (XP) and to re-implement it in Ogre. We also extended the scripting language, leading to the new name OgreVRScript. OgreVRScript allows specifying both the audio and the visual part of an experiment, and the various interactions which are necessary. An important feature of this environment is its extensibility, which allows the programmers to add new features as they are requested for the experimental work. These extensions are automatically available from the scripting language, permitting the continued usage by non-programmers.

OgreAL and OgreVRScript are available on the CROSSMOD SVN.

4.3. LibSL - Simple Library for Graphics

Participant: Sylvain Lefebvre.

Computer Graphics research requires a variety of specialized tools: image and mesh manipulation, interaction with several graphics API on several platforms, graphics processor (GPU) programming.

To ease programming in this context, Sylvain Lefebvre has developed a graphics-programming toolbox, as a C++ library. It simplifies several programming tasks and ease sharing and compilation of code under various environments. This development was started within the EVASION / INRIA Grenoble team and is now continuing within the REVES team. The library is available as a private project on the INRIA Forge.

5. New Results

5.1. Plausible Image Rendering

5.1.1. Antiradiance

Participants: Carsten Dachsbacher, Marc Stamminger, George Drettakis, Frédo Durand.

The creation of realistic computer graphics images involves complex computations to simulate the interaction of light and surfaces. Global illumination algorithms compute physically based and accurate solutions to this problem. Due to the inherent complexity of this task, the computation time is tremendous and as a consequence different approximate algorithms are applied in interactive applications. Typically these algorithms are based on different simplifying assumptions and thus neither convergence with increasing computation time, nor continuous transitions to more accurate methods can be achieved.

In this research project a scalable and flexible global illumination framework has been developed. The key contribution is a reformulation of the process of light transport: instead of computing mutual visibility of surfaces, which is the most time consuming task in all global illumination algorithms, a concept of negative light has been developed which treats visibility implicitly.

By computing global illumination solutions with this implicit visibility we can perform the computations on fast parallel architectures, such as graphics hardware, and adapt the accuracy of the lighting solution to the available computation time: the accuracy of the global illumination ranges from a plausible approximation, achieved at interactive speeds, to a solution as comparable to that obtained from traditional physically based off-line algorithms (Figure 4).

Figure 4. The global illumination for this scene is computed using our Antiradiance method at 10 frames per second, allowing interactive updates when the light source is moved.

This work has been published at SIGGRAPH 2007 [14].

5.1.2. Tile-Trees

Participants: Sylvain Lefebvre, Carsten Dachsbacher.

Texture mapping with atlases suffer from several drawbacks: Wasted memory, seams, uniform resolution and no support of implicit surfaces. Texture mapping in a volume solves most of these issues, but unfortunately it induces an important space and time overhead.

To address this problem, we introduce the Tile-Tree: A novel data structure for texture mapping surfaces. Tile-Trees store square texture tiles into the leaves of an octree surrounding the surface (Figure 5). At rendering time the surface is projected onto the tiles, and the color is retrieved by a simple 2D texture fetch into a tile map. This avoids the difficulties of global planar parameterizations while still mapping large pieces of surface to regular 2D textures. Our method is simple to implement, does not require long pre-processing time, nor any modification of the textured geometry. It is not limited to triangle meshes. The resulting texture has little distortion and is seamlessly interpolated over smooth surfaces. Our method natively supports adaptive resolution.

We show that Tile-Trees are more compact than other volume approaches, while providing fast access to the data. We also describe an interactive painting application, enabling to create, edit and render objects without having to convert between texture representations.

The tiletree project, initiated in 2006, was published at the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D) 2007 Conference where it was presented by Sylvain Lefebvre in April 2007 [19]. The talk was one of the three selected to be presented again at SIGGRAPH 2007 during the 'Reprise from I3D' session. Several aspects of the initial approach were improved, in particular regarding display performance. These improvements have been described in a chapter of the upcoming ShaderX6 book.

5.1.3. Compact Texture Synthesis

Participants: Nader Salman, Sylvain Lefebvre.

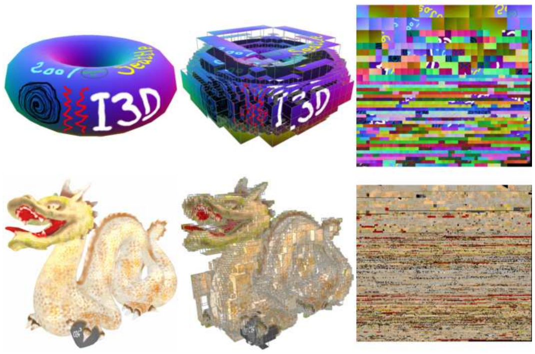

Figure 5. Tile-trees are used to store square texture tiles, permitting compact texturing and interactive editing.

Numerous real-time applications such flight simulators or video games require non-repetitive high-resolution texturing on large landscapes. Such textures may be produced automatically by texture synthesis from example. However, most algorithms output a very large image which must be stored in memory. This project explored a new texture synthesis algorithm which result can be very compactly encoded. The image is only decoded at display time. The encoding scheme is chosen so that the decoder is efficiently implemented on programmable graphics hardware. The approach is based on well known patch based synthesis schemes. In the context of landscape rendering, we also adapt patch positioning to cancel out distortions resulting from the terrain slope. This research was done in the context of the master internship of Nader Salman, and supervised by Sylvain Lefebvre.

5.1.4. Texturing from Photographs

Participants: Christian Eisenacher, Sylvain Lefebvre.

The goal of texture synthesis is to generate an arbitrarily large high-quality texture from a small input sample. Generally, it is assumed that the input image is given as a flat, square piece of texture, thus it has to be carefully prepared from a picture taken under ideal conditions. Instead, in this work we seek to extract the input texture from any surface from within an arbitrary photograph. This introduces several challenges: Only parts of the photograph are covered with the texture of interest, perspective and scene geometry introduce distortions, and the texture is non-uniformly sampled during the capture process. This breaks many of the assumptions used for synthesis. To achieve this goal we combine a simple novel user interface with a generic per-pixel synthesis algorithm to achieve high-quality synthesis from a photograph. Applications range from synthesizing textures directly from photographs to high-quality texture completion. This research was done in the context of the master internship of Christian Eisenacher, and supervised by Sylvain Lefebvre. The work has been conditionally accepted to the EUROGRAPHICS 2008 Conference.

5.1.5. Mesh puzzle

Participants: Sylvain Lefebvre, Carsten Dachsbacher, Marcio Cabral.

A typical approach to quickly create large virtual environments is to rely on basic building-blocks that are combined together to produce more complex layouts. However, creating such building-blocks is extremely difficult and tedious, as they have to enforce strict boundary constraints in order to be combined in a consistent layout. In this ongoing project we explore modeling tools that let the user go through example input 3D datasets and select what he wants to reuse in the final virtual environment. These selected pieces can interactively be added / removed or reshaped to form the final 3D environment. The main contribution is a reshaping operator targeted at architectural pieces (coarsely tessellated geometry and sharp edges) where other kinds of mesh editing approaches, such as Poisson or Laplacian editing, would fail.

5.1.6. *Filtered Tile Maps*

Participant: Sylvain Lefebvre.

One way to obtain very large non-repeating textures is to combine a small set of square textures in a regular grid. By enforcing appropriate boundary constraints the large texture appears seamless. However, this approach has drawbacks when it comes to real time display. In particular, the resulting large texture is improperly filtered and artifacts appear through bilinear interpolation and mipmapping. In this article to appear in the upcoming ShaderX 6 book, Sylvain Lefebvre describes how to properly display such textures using latest programmable graphics hardware.

5.1.7. *Compressed Random Accessed Trees for Spatially Coherent Data*

Participants: Sylvain Lefebvre, Hugues Hoppe.

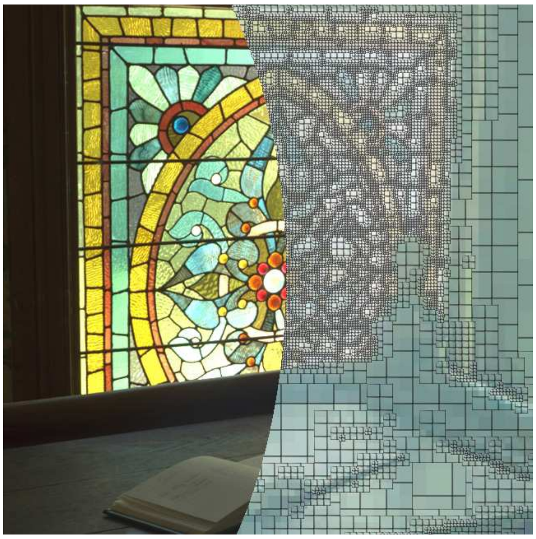

Figure 6. A compressed random access tree capturing the large scale variations of a high-dynamic range image. This floating point RGB image compresses at 5 bits/pixel using our approach.

Adaptive multiresolution hierarchies are highly efficient at representing spatially coherent graphics data. This research introduces a framework for compressing such adaptive hierarchies using a compact randomly-accessible tree structure. Prior schemes have explored compressed trees, but nearly all involve entropy coding of a sequential traversal, thus preventing fine-grain random queries required by rendering algorithms. Instead, fixed-rate encoding is used for both the tree topology and its data. Key elements include the replacement of pointers by local offsets, a forested mipmap structure, vector quantization of inter-level residuals, and efficient coding of partially defined data. Both the offsets and codebook indices are stored as byte records for easy parsing by either CPU or GPU shaders. We show that continuous mipmapping over an adaptive tree is more efficient using primal subdivision than traditional dual subdivision. Finally, efficient compression is demonstrated on many data types including light maps, alpha mattes, distance fields, and HDR images (see Figure 6). This work is collaboration between Sylvain Lefebvre and Hugues Hoppe (Microsoft Research USA). It was published at the 2007 Eurographics Symposium on Rendering where it was presented by Sylvain Lefebvre [20].

5.1.8. *Interactive image editing*

Participants: Sylvain Lefebvre, Olga Sorkine, Mathias Eitz, Andrew Nealen, Marc Alexa.

This ongoing project explores an interactive image editing tool that allows the user to move or distort objects within the image using a simple click-and-drag interface. The key idea is to move the object - or parts of the object - by generating seams around it, such that some seams contract and others expand, thus translating the area encircled by the seams. In particular we focus on the detections of areas where completion will be easier due to specific properties of the image content (stochastic, smooth, ...) This work is collaboration between Sylvain Lefebvre, Olga Sorkine (TU. Berlin), Mathias Eitz (TU Berlin), Andrew Nealen (TU Berlin) and Marc Alexa (TU Berlin).

5.1.9. Surface Solid Texture Synthesis from Slices

Participants: Sylvain Lefebvre, George Drettakis, Xin Tong, Yue Dong.

This ongoing project explores a solid texture synthesis algorithm that focuses computations only around the surface to be textured. In contrast to previous solid texture synthesis approaches, it avoids the wasteful storage of the entire volume. The result is the same as if the surface was carved out of a full volume. The key contribution is a new parallel, deterministic, volumetric synthesis scheme. This work is a collaboration between Sylvain Lefebvre, George Drettakis, Xin Tong (Microsoft Research Asia) and Yue Dong (Microsoft Research Asia).

5.1.10. An Interactive Perceptual Rendering Pipeline using Contrast and Spatial Masking

Participants: George Drettakis, Nicolas Bonneel, Sylvain Lefebvre, Carsten Dachsbacher, Michael Schwarz, Isabelle Viaud-Delmon.

Figure 7. The rendering of a complex museum scene with our pipeline calibrated for a 512x512 Resolution image (left) and with a distance based LOD selection (right). The zoomed insets clearly show the difference in the choice of LOD

We developed a pipeline which allows the rendering of highly complex scenes taking into account perceptual phenomena such as contrast and spatial masking which are due to the interaction of one or a set of objects with other objects (Figure 7).

Our first contribution is the developpment of a GPU-based perceptual rendering framework. To integre prediction of masking into a GPU pipeline, we chose to split the scene into layers, allowing us to take into account inter-object masking. Layer rendering and appropriate combinations all occur on the GPU, and are followed by the efficient computation of a threshold map, again on the graphics processor. As a consequence, no CPU-readback is required, and the framework is very efficient and adapted to the modern rendering pipeline, allowing prediction of inter-object visual masking at interactive rates.

We then perform a level-of-detail (LOD) selection strategy based on the layer and the predicted contrast and spatial masking for each object. This is achieved using a fragment program and occlusion queries, which are the fastest GPU to CPU feedback mechanism; as a result, the overhead for this step is minimal.

We performed a perceptual user study to validate our approach. This consisted in showing pairs of scenes, one containing an animated gargoyle whose LOD has been chosen by our algorithm, and the other containing the gorgoyle at a different LOD (better quality, lower quality, or highest possible quality), as predicted by our GPU pipeline. We then ask the user if she perceives a difference between the two images or not.

The results indicate that our algorithmic choices are consistent with the perceived differences in images.

This work was published at the 2007 Eurographics Symposium on Rendering [17].

5.2. Plausible Audio Rendering

5.2.1. 3D Audio matting and re-rendering

Participants: Nicolas Tsingos, Emmanuel Gallo.

We proposed a novel approach to real-time spatial rendering of realistic auditory environments and sound sources recorded live, in the field (Figure 8). Using a set of standard microphones distributed throughout a real-world environment we record the sound-field simultaneously from several locations. After spatial calibration, we segment from this set of recordings a number of auditory components, together with their location. We compared existing time-delay of arrival estimations techniques between pairs of widely-spaced microphones and introduced a novel efficient hierarchical localization algorithm. Using the high-level representation thus obtained, we can edit and re-render the acquired auditory scene over a variety of listening setups. In particular, we can move or alter the different sound sources and arbitrarily choose the listening position. We can also composite elements of different scenes together in a spatially consistent way.

Our approach provides efficient rendering of complex soundscapes which would be challenging to model using discrete point sources and traditional virtual acoustics techniques.

This works was published in a special issue of the EURASIP Journal on Advances in Signal Processing [16].

We further improved our segmentation strategy by separating foreground, well-localized sounds and more diffuse background sounds and conducted perceptual tests to compare our approach to existing alternatives. The results were published at the 30th Audio Engineering Society International Conference on Intelligent Audio Environments [18].

5.2.2. Instant sound scattering

Participants: Nicolas Tsingos, Carsten Dachsbacher, Sylvain Lefebvre, Matteo Dellepiane.

Real-time sound rendering engines often render occlusion and early sound reflection effects using geometrical techniques such as ray or beam tracing. They can only achieve interactive rendering for environments of low local complexity resulting in crude effects which can degrade the sense of immersion. However, surface detail or complex dynamic geometry has a strong influence on sound propagation and the resulting auditory perception.

This project focuses on high-quality modeling of first-order sound scattering. Based on a surface-integral formulation and the Kirchhoff approximation, we propose an efficient evaluation of scattering effects, including both diffraction and reflection, that leverages programmable graphics hardware for dense sampling of complex surfaces. We evaluate possible surface simplification techniques and show that combined normal and displacement maps can be successfully used for audio scattering calculations.

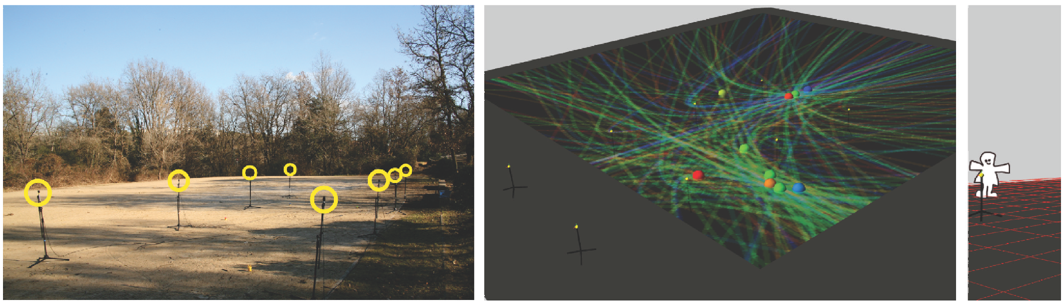

Figure 8. Left: We use multiple arbitrarily positioned microphones (circled in yellow) to simultaneously record real-life auditory environments. Middle: We analyze the recordings to extract the positions of various sound components through time. Right: This high-level representation allows for post-editing and re-rendering the acquired soundscape within generic 3D audio rendering architectures.

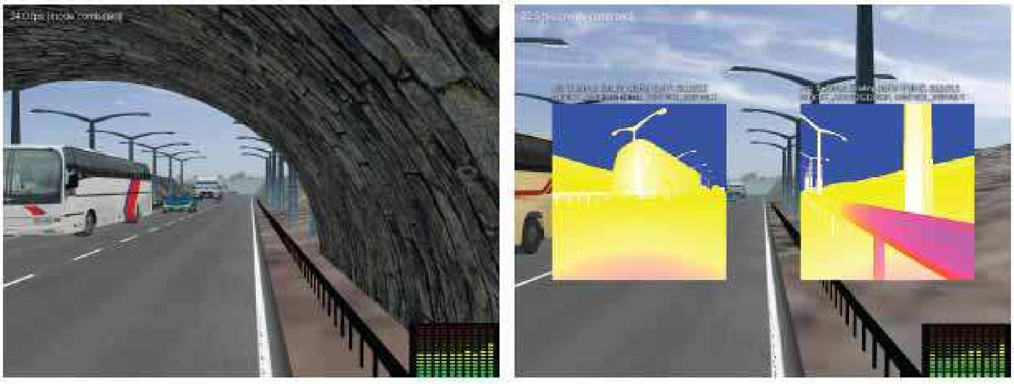

Figure 9. Our approach applied to an interactive driving simulation. Sounds from the cars are scattered by the surrounding obstacles, e.g. a tunnel (left). Two views of the obstacles in the environment used to compute the scattering integral are shown in the right-hand side image.

Contrary to traditional geometrical acoustics approaches, our framework simulates sound scattering by evaluating an integral on all visible surfaces from the sound sources. This allows us to capture effects closer to wave acoustics without performing a full boundary element simulation. However, a direct implementation of such an approach would be extremely time-consumming. In this project, we cast this problem in a framework based on graphics processors (GPU) for all geometric computations (Figure 9).

There are three main advantages to this usage of the GPU. First, the raw power of rasterization allows the determination of visible surfaces from sound sources at speeds orders of magnitude faster than previous methods, thanks to the massively parallel pipeline. Any type of primitive that can be rasterized can be used with our approach, for instance direct point-based representations from scanned data. Second, the ability to use a mip-mapping strategy, allows us to compute the scattering integral extremely efficiently. Third, the ability to use effective level of detail mechanisms such as displacement mapping, allows us to treat extremely complex geometry as textured billboards.

For instance, this technique was applied on several complex geometries (for example, a 10M triangles model of the façade of Pisa Cathedral, obtained using a time-of -flight scanner), and compared with recordings made on-site. We also applied it to the Kulkulkan temple, a Maya pyramid exhibiting very audible scattering effects due to diffraction on its staircase. Our approach convincingly captures this diffraction effect and the corresponding chirped echo.

This work was published at the 2007 Eurographics Symposium on Rendering [25].

5.2.3. Reverberation graphs

Participants: Efstathios Stavrakis, Nicolas Tsingos.

Proper modeling of sound propagation is very important for virtual acoustics and virtual reality applications. In particular, reverberation effects due to sound reflections off wall surfaces carry major cues related the size of the environment and distance to sound sources. As a result, reverberation is arguably one of the most important audio effects to simulate in virtual environments. In this project, we concentrate on interactive, yet physically-based, modeling of reverberation effects in complex architectural environments containing arbitrary coupled acoustic spaces. Our approach combines characteristics of geometrical and statistical acoustics to achieve both low latency and good audio quality. Akin to geometrical acoustics simulations, it requires a 3D input model. From this model, a spatial adjacency graph between the coupled acoustic spaces is constructed. Energy decay profiles are then determined *off-line* for each space using either path tracing simulations or simpler statistical methods. The complex global coupling of all the spaces is solved *on-line* by exploring all energy propagation routes along the spatial adjacency graph.

5.2.4. Processing Massive Impact Sounds

Participants: Nicolas Bonneel, Nicolas Tsingos, George Drettakis, Isabelle Viaud-Delmon.

There are often occasions when a massive number of impact sounds occurs: the impact of debris in an explosion, large number of objects being unloaded off a truck, objects being blown away in a strong wind. Processing these impact sounds is hard, due to their sheer number and the expensive of synthesizing the sounds. In the context of the Ph.D. of N. Bonneel, we are developing a novel solution allowing the integration of modal synthesis for impact sounds into the perceptual processing pipeline (see [21]). This involves a fast approximation of signal energy without performing the modal synthesis and a fast approach to transform modes into the Fourier domain. This work is planned for submission early 2008.

5.3. Virtual Environments with Sound

5.3.1. Progressive Perceptual Audio Rendering of Complex Scenes

Participants: Nicolas Tsingos, Thomas Moeck, Nicolas Bonneel.

Despite recent advances, including sound source clustering and perceptual auditory masking, high quality rendering of complex virtual scenes with thousands of sound sources remains a challenge. Two major bottlenecks appear as the scene complexity increases: the cost of clustering itself, and the cost of pre-mixing source signals within each cluster.

In this work, we first propose an improved hierarchical clustering algorithm that remains efficient for large numbers of sources and clusters while providing progressive refinement capabilities. We have also developed a lossy pre-mixing method based on a progressive representation of the input audio signals and the perceptual importance of each sound source. We performed quality evaluation user tests which indicate that the recently introduced audio saliency map is inappropriate for this task. Consequently we propose a “pinnacle", loudness-based metric, which gives the best results for a variety of target computing budgets. We also performed a perceptual pilot study which indicates that in audio-visual environments, it is better to allocate more clusters to visible sound sources (Figure 10). We propose a new clustering metric using this result. As a result of these three solutions, our system can provide high quality rendering of thousands of 3D-sound sources on a “gamer-style" PC.

Figure 10. Left: sound source clustering without the audio-visual metric. Right: the clusters with our new metric. The new metric better separates the sources in the visual field.

This work was published at ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games 2007 (I3D) [21].

5.3.2. *Contact sound shading*

Participants: Cécile Picard, Nicolas Tsingos.

The goal is the real-time synthesis of sounds resulting from physical interaction of various objects in a 3D virtual environment (e.g. collisions, sliding, rolling, etc.) which are difficult to create in a pre-production process since they are highly dynamic and vary drastically depending on objects and the interaction type. In particular in the context of actual video games, the solid dynamics simulation generally use simplified models of the scene that cannot directly be used to synthesize compelling contact sounds. The study focus on providing several detail-layer mechanism that allow for realistic contact sound modeling while maintaining simple solid dynamic simulations. The purpose is to improve the quality and the authoring aspect of nowadays real-time synthesis of complex contact sounds. More precisely, the study aims at providing improvements to vibrational data synthesis, contact/interaction force synthesis and detail sound texture synthesis. This approach may unify and develop previous work in the same framework.

5.3.3. *Auditory Masking for Scalable VOIP Bridges*

Participants: Nicolas Tsingos, Arnault Nagle, Guillaume Lemaitre, Aurelien Sollaud.

Endpoints or conference servers of current audio-conferencing solutions use all the audio frames they receive in order to mix them into one final aggregate stream. However, at each time-instant, some of this content may not be audible due to auditory masking. Hence, sending corresponding frames through the network leads to a loss of bandwidth, while decoding them for mixing or spatial audio processing leads to increased processor load. In this project, we developped a solution based on an efficient on-the-fly auditory masking evaluation. Our technique allows prioritizing audio frames in order to select only those audible for each connected client. Results of quality tests show that our approach is perceptually transparent. We integrated it in a France Telecom audio conference server. Tests in a 3D game (Figure 11) environment with spatialized chat capabilities show a 70% average reduction in required bandwidth.

This work was published [22] at the 30th Audio Engineering Society International Conference on Intelligent Audio Environments (see http://www-sop.inria.fr/reves/OPERA for additional info).

Figure 11. The video game "Flower Power Shooter" from the company Virtools was augmented with on-line 3D chat capabilities. The voice over IP servers used for streaming the speech data implements on-the-fly auditory masking evaluation from the point of view of each partipant in order to reduce the bandwidth.

5.3.4. *Design and Evaluation of a Real-World Virtual Environment for Architecture and Urban Planning*

Participants: George Drettakis, Maria Roussou, Nicolas Tsingos, Alex Reche.

Figure 12. Left: A snapshot of the virtual environment with the tramway simulation. Right: An architect using our VE system for evaluation of his design.

In this project we investigate a user-centered design approach to the development of a Virtual Environment (VE), by utilizing an iterative, user-informed process throughout the entire design and development cycle. A preliminary survey was first undertaken with endusers, i.e., architects, chief engineers and decision makers of a real-world architectural and urban planning project, followed by a study of the traditional workflow employed. We then determined the elements required to make the VE useful in the real-world setting, choosing appropriate graphical and auditory techniques to develop audiovisual VEs with a high level of realism (Figure 12). Our user-centered design approach guided the development of an appropriate interface and an evaluation methodology to test the overall usability of the system. The VE was evaluated both in the laboratory and, most importantly, in the users natural work environments. In this study we present the choices we made as part of the design and evaluation methodologies employed, which successfully combined research goals with those of a real-world project.

Among other results, this evaluation suggests that involving users and designers from the beginning improves the effectiveness of the VE in the context of the real world urban planning project. Furthermore, it demonstrates that appropriate levels of realism, in particular spatialized 3D sound, high-detail vegetation and shadows, as well as the presence of rendered crowds, are significant for the design process and for communicating about designs; they respectively enable better appreciation of overall ambience of the VE, perception of space and physical objects as well as the sense of scale. We believe this study is of interest to VE researchers, designers and practitioners, as well as professionals interested in using VR in their workplace.

This work was performed as part of our FP5 EU project CREATE, in collaboration with the “Mission Tramway" in Nice, and was published in the journal “Presence: Teleoperators and Virtual Environments" [15].

5.3.5. *Role of semantic vs spatial congruency in a bimodal go/no-go task*

Participants: Nicolas Bonneel, Clara Suied, Isabelle Viaud-Delmon.

In the context of the European project CROSSMOD, Nicolas Bonneel participated in a neuroscience experiment concerning the role of semantic vs spatial congruency in a bimodal go/no-go task. Numerous factors, both bottom-up and cognitive, have been shown to contribute to the integration of visual and auditory stimuli. It has been clearly demonstrated that the spatial relationship between the visual and auditory stimuli influence behavioural performances such as reaction times. Moreover, the role of semantic congruence on the enhancement of performance has also been highlighted. In realistic virtual environments, both spatial and semantic characteristics might be at stake for optimal bimodal integration. The aim of this study is to compare the role of spatial and semantic congruencies on behavioral performances (reaction times) during immersion in virtual reality.

The conclusions of the experiment were that responses were found to be faster in the presence of semantic congruence (i.e., the sound and image correspond), while there was no significant effect of spatial misalignment in reaction times. This indicates dominance of semantic content over spatial information in bimodal (audio-visual) integration. This work was presented at the International Multisensory Research Forum [23].

6. Contracts and Grants with Industry

6.1. Alias|Wavefront

We are part of the Alias|Wavefront software donation program. We use Maya extensively to model various objects used in our animations and results, for many of our projects.

6.2. Eden Games

We have an ongoing collaboration with Eden Games (an ATARI game studio in Lyon, France). Eden had previously licensed our audio masking/clustering technology for TestDrive Unlimited and Alone in the Dark Near Death Experience on Xbox360, PCs and PS3. We extended our licensing agreement to all platforms and two new titles. The new agreement will be used to fund a PhD thesis in order co-develop novel real-time synthesis algorithms for interactive audio rendering. This PhD started January 2007.

6.3. CIFRE CSTB

Participants: George Drettakis, Nicolas Tsingos, Emmanuel Gallo.

We held a CIFRE contract with the CSTB for the thesis of Emmanuel Gallo. He works on sound rendering for urban simulation, in the context of the OPERA project. In this context, Emmanuel Gallo spends a significant amount of time at the CSTB and strives to apply his work directly on urban planning and architectural applications.

6.4. Renault

We have signed a collaboration agreement with Renault for evaluation of the audio matting technique. Renault also provided a set of measurements which can be processed by our approach.

7. Other Grants and Activities

7.1. Regional/Local Projects

7.1.1. Collaboration with CNRS and IRCAM

In the context of the CROSSMOD European project we have developed continuous and active collaboration with Isabelle Viaud-Delmon of CNRS (Laboratoire Vulnérabilité, adaptation et psychopathologie, UMR 7593) and Olivier Warusfel of IRCAM. The collaboration with CNRS has helped us in the development of experimental protocols for perceptual validation work, and has resulted in a common paper submssion. Similarly, our collaboration with IRCAM has been very beneficial for our work in acoustics.

7.1.2. Collaboration with CSTB Sophia-Antipolis

Participants: George Drettakis, Nicolas Tsingos, Emmanuel Gallo.

The collaboration with CSTB continued through the PhD research of Emmanuel Gallo (CIFRE fellowship program).

7.1.3. The workbench platform

Participants: David Geldreich, George Drettakis, Nicolas Tsingos.

The regional Provence-Alpes-Cote d'Azur government has co-funded (with INRIA) the acquisition of semi-immersive platform for virtual environments, also known as "workbench". David Geldreich initially setup the system, and now continues the support.

The platform is composed of a Barco Baron screen (1.5 m diagonal) which can be tilted from near horizontal (table) to near vertical position. The screen is equiped with a BarcoReality 908 CRT projector driven by an off-the-shelf PC (2 Intel Xeon 2.8GHz + GeForce 6800 Ultra AGP 8x graphics) running under Linux and Windows XP. Stereo display is achieved through a frequency-doubler StereoGraphics EPC-2 and active LCD shutter-glasses (StereoGraphics CrystalEyes and NuVision/60GX). Finally, we also use a 6-DOF Polhemus Fastrak 3D tracking system interfaced with a stylus and a gamepad for interaction, and an additional captor for view-point tracking and view-dependent rendering.

For the audio, we use six Yamaha MSP3 speakers driven by a Motu 896 firewire audio interface.

D. Geldreich installed the system and developed a suite of APIs and tools allowing for easy integration of the stereo display and tracking system in any Performer/VTK- based application. We have recently upgraded our tracking facility to an ART, vision-based system.

This system is a shared ressource used by both our group and others. We used it both as a demo platform and an immersive display used for conducting experiments and user testing.

7.2. European Projects

7.2.1. Open FET IST CROSSMOD

Participants: George Drettakis, Monique Varlet, David Geldreich, Nicolas Tsingos, Nicolas Bonneel, Sylvain Lefebvre, Isabelle Viaud-Delmon, Carsten Dachsbacher, Thomas Moeck, Efstathios Stavrakis.

CROSSMOD, "Cross-modal Perceptual Interaction and Rendering", is an Open FET (Future and Emerging Technologies) STREPS of the 6th FP IST. REVES coordinates this project; the other partners are CNRS, IRCAM (FR), University of Bristol (UK), Vienna University of Technology (A) University of Erlangen (D) and the CNR in Pisa (I).

CROSSMOD has three main objectives: the advancement of fundamental understanding of cross-modal effects for the display of virtual environments (VEs), the development of novel algorithms based on this understanding, and their evaluation using target applications. We concurrently develop the fundamental and foundational aspects related to audiovisual perception and display, as well as the novel and concrete algorithmic solutions. The scope of CROSSMOD is thus quite fundamental; however, its success will be measured by the efficiency/quality improvements of the new algorithmic solutions developed, and the validation of our hypotheses in realistic VE settings.

In the second year of CROSSMOD, REVES had a number of significant research activities. These activities are described in more detail in the appropriate sections of this report. The published research results concern the work on crossmodal audio clustering [21], the work on GPU-based threshold maps [17], the Instant Sound Scattering [25] method and the semantic vs. spatial congruency experiment [23]. Other CROSSMOD-related work includes investigation of the effect of sound on material recognition, and the work on massive impact sounds.

As project coordinator, we are responsible for the publication of all the internal wiki material and continued updates, and the external website. We also have the overall responsibility for all project Deliverables and in particular the 6-month and annual reports. As coordinator, REVES also is responsible for the overall coordination of the scientific and management activity of CROSSMOD. Given its exploratory and multidisciplinary nature, this is a challenging, but enriching, task.

7.3. Visiting Researchers

Christian Eisenacher, from the University of Erlangen, completed his masters internship at REVES (November 2006-April 2007). His work resulted in a conditionally accepted publication (EG08).

Isabelle Viaud-Delmon and Clara Suied from CNRS visited for 2 days in August in the context of the CROSSMOD project.

We hosted several visiting researchers this year: Dinesh Pai (University of British Columbia), Hugues Hoppe (Microsoft Research US), Olga Sorkine and Andy Nealen (Technical University of Berlin), Sylvain Paris (MIT, now at Adobe US), Efstathios Stavrakis (then at VUT) and Kaleigh Smith (MPII).

We had four visitors before and after the Eurographics programme committee meeting in Nice in November: Doug James (Cornell), Frédo Durand (MIT), Bruno Levy (INRIA Nancy) and Eugene Fiume (University of Toronto).

7.4. Bilateral Collaborations

7.4.1. France-China

We collaborate with Microsoft Research Asia (Beijing) on texture synthesis (X. Tong).

7.4.2. France-Germany

Two students from the graphics group of Marc Stamminger at the University of Erlangen did their master thesis at REVES: Thomas Moeck and Christian Eisenacher. Our collaboration with Erlangen has resulted in joint publications at SIGGRAPH and I3D and an accepted publication at Eurographics 2008.

We are also collaborating with the University of Stuttgart (new position of C. Dachsbacher), the University of Konstanz (O. Deussen) and the Technical University of Berlin (O. Sorkine).

7.4.3. France-United States of America

The collaboration with Frédo Durand of MIT CSAIL has resulted in a publication this year at SIGGRAPH.

We also have an on-going collaboration with Hugues Hoppe from Microsoft Research, which resulted in a publication at EGSR.

7.4.4. France-Italy

In the context of the CROSSMOD project, we have developed a working relationship with CNR Pisa; this resulted in a joint publication at EGSR with Matteo Dellepiane.

8. Dissemination

8.1. Participation in the Community

8.1.1. Program Committees

George Drettakis was co-chair of the programme committee Eurographics 2008, which will take place in Crete in April 2008. This involved organising the entire review process for the conference (300 submissions), and the committee meeting (52 expert members of the community) in Nice. Catherine Martin (the assistant of the ORION project-team) assisted in this organisation. G. Drettakis was also a member of the programme committee Eurographics 2008 and for EGSR 2007 and 2008. He served as a reviewer for SIGGRAPH 2008 and various journals.

Nicolas Tsingos was program committee member of EGSR 2007/2008, ACM MM art program 2007, SPM 2007/2008 and Pacific Graphics 2007. In 2007 he was also reviewer for SIGGRAPH, EGSR, EURASIP JASP, Pacific Graphics, ACM Trans. on Applied Perception, I3D, IEEE Trans. Vis. and Comp. Graphics, EUROGRAPHICS and SPM.

Sylvain Lefebvre was reviewer for the TVCG and TOG journals, SIGGRAPH 2006, GI 2006 and Pacific Graphics 2006. In 2007, Sylvain Lefebvre was part of the Eurographics 2007 Committee, the Eurographics 2008 Committee, the I3D 2008 Committee and the video track committee of the Symposium on Computational Geometry 2007. He was reviewer for the SIGGRAPH 2007 Conference, the 2007 Eurographics Symposium on Rendering and the Computer Graphics Forum journal.
Carsten Dachsbacher was a member of the SIBGRAPI 2006, GRAPP'07 and Shader X5 committees. He was reviewer for I3D 2007, SIGGRAPH 2006, Eurographics 2006, EGSR 2006, SIBGRAPI 2006, Computer Graphics Forum, Graphical Models, Pacific Graphics 2006, IEEE Vis 2006, and Shader X5.

8.1.2. Distinctions

G. Drettakis won the Eurographics "Outstanding Technical Achievement Award" 2007 and was named Fellow of the Eurographics Association.

8.1.3. Invited Talks

G. Drettakis presented a talk on Audiovisual Perceptual Rendering at ETH Zurich (June).

8.1.4. Thesis Committees

G. Drettakis was an external examiner or committee member for the Ph.D. theses of C. Bosch (University of Girona) and B. Segovia (University of Lyon) and the Habilitation thesis of N. Holzscuch (INRIA Grenoble).

8.1.5. COST and CUMIR

George Drettakis was the coordinator of the Working Group on International Relations of the INRIA national level COST (Scientific and Technological Orientation Coucil/Conseil d'Orientation Scientifique et Technologique) until September this year. The specific working group is responsible for the evaluation of all INRIA-supported international activities, including the International Associated Team programme, and the various bi-lateral and regional international cooperations. The group also makes proposals to the INRIA national directorate on issues relating to general international policy. This activity involves the coordination and organisation of these evaluations, the coordination of a group of 6-8 researchers, one annual meeting of the group, and participation in the bi-monthly meetings of the COST.

Nicolas Tsingos is a member of the CUMIR, the researcher users group, serving as an interface between the researchers of the Sophia-Antipolis INRIA Unit and the computer services. This year, he has been responsible for several activities, including a poster for the CUMIR and the project to determine the renewal of hardware on the site.

8.1.6. Web server

Participants: George Drettakis, Sylvain Lefebvre.

http://www-sop.inria.fr/reves/

The project web-server is constantly updated with our research results. Most of the publications of REVES can be found online, and often include short movies demonstrating our research results. See http://www-sop.inria.fr/reves/publications/index.php3?LANG=gb

Sylvain Lefebvre installed the bibliography server Basilic (from Grenoble), which allows a much improved presentation of our publications.

8.2. Teaching

8.2.1. University teaching

George Drettakis was responsible for the Computer Graphics course at ISIA (Ecole des mines) in 2007, given in May and September (15 hours), with the participation of Nicolas Tsingos (10 hours) and S. Lefebvre. He also taught a Computer Graphics course at Ecole Centrale de Paris with Sylvain Lefebvre and Nicolas Tsingos (6 hours each for a total of 18 hours).

8.2.2. PhD Thesis Completed and Continuing

Emmanuel Gallo defended his Ph.D. in March 2007. He is now working for the company *RealViz*. Cécile Picard started her PhD in January 2007 and is working on physically-based audio synthesis. Nicolas Bonneel started his PhD in August 2006 as part of the CROSSMOD project is working on audiovisual crossmodal effects. Marcio Cabral started his thesis in September 2007 and is working on procedural techniques for modelling and on image-based relighting.

8.3. Participation at conferences

8.3.1. Presentations at Conferences

Sylvain Lefebvre attended the I3D 2007 conference where he presented the work on TileTrees. He was one of the presenters of the SIGGRAPH 2007 course on example based texture synthesis, together with Vivek Kwatra (University of North California at Chapel Hill), Li-Yi Wei (Microsoft Research Asia) and Greg Turk (Georgia Institute of Technology). The TileTrees talk was also selected to be presented again at SIGGRAPH in the 'Reprise from I3D' session.

Nicolas Tsingos gave two invited talks at the Intl. Congress on Acoustics 2007, held in Madrid, Spain.

G. Drettakis presented the spatial masking paper, N. Tsingos the Instant Sound paper and S. Lefebvre the compressed random access trees paper at EGSR07 in Grenoble (June).

8.3.2. Participation at Conferences and Workshops

N. Tsingos went to the AES convention in May in Vienna and to the ICA2007 and SIGGRAPH 2007 conferences. He also attended the SIGGRAPH 2007 conference where he participated in the course on examples based texture synthesis. N. Bonneel and C. Dachsbacher attended EGSR 07 in Grenoble.

8.4. Demonstrations and Press

8.4.1. Demonstrations

Participants: David Geldreich, George Drettakis, Nicolas Tsingos.

We presented demonstrations of our research to representatives of the Conseil Regional PACA (February), representatives of the Conseil General des Alpes Marities (April), representatives of the Office du Tourisme de Nice (May), M. Ghallab (Scientific Director of INRIA), students of the Ecole Centrale de Paris (October), and the Technicom company (October). N. Tsingos will be presenting a demonstration of our Audile audio technology for the 40 years of INRIA celebration in December.

8.4.2. Press

Nicolas Tsingos was interviewed for the journal *Physics Today* in the context of a feature paper on physically-based animation and sound synthesis. David Geldreich was interviewed for an article in a popular magazine on Computers, 01.net on Ogre3D. G. Drettakis participated in a live radio programme ("Science Publique") on the France Culture station in November.

9. Bibliography

Major publications by the team in recent years

[1] O. DEUSSEN, C. COLDITZ, M. STAMMINGER, G. DRETTAKIS. *Interactive visualization of complex plant ecosystems*, in "Proceedings of the IEEE Visualization Conference", IEEE, October 2002, http://www-sop.inria.fr/reves/publications/data/2002/DCSD02.

[2] F. DUGUET, G. DRETTAKIS. *Robust Epsilon Visibility*, in "Proceedings of ACM SIGGRAPH 2002", J. HUGHES (editor), ACM Press / ACM SIGGRAPH, July 2002, http://www-sop.inria.fr/reves/publications/data/2002/DD02.

[3] F. DUGUET, G. DRETTAKIS. *Flexible Point-Based Rendering on Mobile Devices*, in "IEEE Computer Graphics and Applications", vol. 24, n^o 4, July-August 2004, http://www-sop.inria.fr/reves/publications/data/2004/DD04.

[4] P. ESTALELLA, I. MARTIN, G. DRETTAKIS, D. TOST. *A GPU-driven Algorithm for Accurate Interactive Reflections on Curved Objects*, in "Rendering Techniques (Proceedings of the Eurographics Symposium on Rendering)", T. AKENINE-M" OLLER, W. HEIDRICH (editors), Eurographics/ACM SIGGRAPH, June 2006, http://www-sop.inria.fr/reves/Basilic/2006/EMDT06.

[5] T. FUNKHOUSER, N. TSINGOS, I. CARLBOM, G. ELKO, M. SONDHI, J. WEST, G. PINGALI, P. MIN, A. NGAN. *A Beam Tracing Method for Interactive Architectural Acoustics*, in "The Journal of the Acoustical Society of America (JASA)", vol. 115, n^o 2, February 2004, http://www-sop.inria.fr/reves/Basilic/2003/FTCESWPMN03.

[6] X. GRANIER, G. DRETTAKIS. *A Final Reconstruction Framework for an Unified Global Illumination Algorithm*, in "ACM Transactions on Graphics", vol. 23, n^o 2, April 2004, http://www-sop.inria.fr/reves/publications/data/2004/GD04.

[7] P. POULIN, M. STAMMINGER, F. DURANLEAU, M.-C. FRASSON, G. DRETTAKIS. *Interactive Point-based Modeling of Complex Objects from Images*, in "Proceedings of Graphics Interface", June 2003, http://www-sop.inria.fr/reves/Basilic/2003/PSDFD03.

[8] A. RECHE, I. MARTIN, G. DRETTAKIS. *Volumetric Reconstruction and Interactive Rendering of Trees from Photographs*, in "ACM Transactions on Graphics (SIGGRAPH Conference Proceedings)", vol. 23, n^o 3, July 2004, http://www-sop.inria.fr/reves/publications/data/2004/RMD04.

[9] M. STAMMINGER, G. DRETTAKIS. *Interactive Sampling and Rendering for Complex and Procedural Geometry*, in "Rendering Techniques (Proceedings of the Eurographics Workshop on Rendering)", K. MYSKOWSKI, S. GORTLER (editors), 12th Eurographics workshop on Rendering, Springer Verlag, Eurographics, 2001, http://www-sop.inria.fr/reves/Basilic/2001/SD01.

[10] M. STAMMINGER, G. DRETTAKIS. *Perspective Shadow Maps*, in "Proceedings of ACM SIGGRAPH 2002", J. HUGHES (editor), ACM Press/ ACM SIGGRAPH, July 2002, http://www-sop.inria.fr/reves/publications/data/2002/SD02.

[11] N. TSINGOS, I. CARLBOM, G. ELKO, R. KUBLI, T. FUNKHOUSER. *Validating acoustical simulations in the Bell Labs Box*, in "Computer Graphics and Applications, IEEE", vol. 22, n^o 4, Jul/Aug 2002, p. 28-37.

[12] N. TSINGOS, E. GALLO, G. DRETTAKIS. *Perceptual Audio Rendering of Complex Virtual Environments*, in "ACM Transactions on Graphics (SIGGRAPH Conference Proceedings)", vol. 23, n^o 3, July 2004, http://www-sop.inria.fr/reves/publications/data/2004/TGD04.

[13] B. WALTER, G. DRETTAKIS, D. GREENBERG. *Enhancing and Optimizing the Render Cache*, in "Proceedings of the Eurographics Workshop on Rendering", P. DEBEVEC, S. GIBSON (editors), ACM Press, Eurographics, June 2002, http://www-sop.inria.fr/reves/Basilic/2002/WDG02.

Year Publications

Articles in refereed journals and book chapters

[14] C. DACHSBACHER, M. STAMMINGER, G. DRETTAKIS, F. DURAND. *Implicit Visibility and Antiradiance for Interactive Global Illumination*, in "ACM Transactions on Graphics (SIGGRAPH Conference Proceedings)", vol. 26, n^{o} 3, August 2007, http://www-sop.inria.fr/reves/Basilic/2007/DSDD07.

[15] G. DRETTAKIS, M. ROUSSOU, A. RECHE, N. TSINGOS. *Design and Evaluation of a Real-World Virtual Environment for Architecture and Urban Planning*, in "Presence: Teleoperators & Virtual Environments, MIT Press", 2007, http://www-sop.inria.fr/reves/Basilic/2007/DRRT07.

[16] E. GALLO, N. TSINGOS, G. LEMAITRE. *3D-Audio Matting, Post-editing and Re-rendering from Field Recordings*, in "EURASIP: Journal on Advances in Signal Processing", Special issue on Spatial Sound and Virtual Acoustics, 2007, http://www-sop.inria.fr/reves/Basilic/2007/GTL07.

Publications in Conferences and Workshops

[17] G. DRETTAKIS, N. BONNEEL, C. DACHSBACHER, S. LEFEBVRE, M. SCHWARZ, I. VIAUD-DELMON. *An Interactive Perceptual Rendering Pipeline using Contrast and Spatial Masking*, in "Rendering Techniques (Proceedings of the Eurographics Symposium on Rendering)", Eurographics, June 2007, http://www-sop.inria.fr/reves/Basilic/2007/DBDLSV07.

[18] E. GALLO, N. TSINGOS. *Extracting and Re-rendering Structured Auditory Scenes from Field Recordings*, in "AES 30th International Conference on Intelligent Audio Environments", 2007, http://www-sop.inria.fr/reves/Basilic/2007/GT07.

[19] S. LEFEBVRE, C. DACHSBACHER. *TileTrees*, in "Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games", ACM Press, ACM SIGGRAPH, 2007, http://www-sop.inria.fr/reves/Basilic/2007/LD07.

[20] S. LEFEBVRE, H. HOPPE. *Compressed Random-Access Trees for Spatially Coherent Data*, in "Rendering Techniques (Proceedings of the Eurographics Symposium on Rendering)", Eurographics, 2007, http://www-sop.inria.fr/reves/Basilic/2007/LH07.

[21] T. MOECK, N. BONNEEL, N. TSINGOS, G. DRETTAKIS, I. VIAUD-DELMON, D. ALOZA. *Progressive Perceptual Audio Rendering of Complex Scenes*, in "Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games", ACM SIGGRAPH, April 2007, http://www-sop.inria.fr/reves/Basilic/2007/MBTDVA07.

[22] A. NAGLE, N. TSINGOS, G. LEMAITRE, A. SOLLAUD. *On-the-fly Auditory Masking for Scalable VOIP Bridges*, in "AES 30th International Conference on Intelligent Audio Environments", 2007, http://www-sop.inria.fr/reves/Basilic/2007/NTLS07.

[23] C. SUIED, N. BONNEEL, I. VIAUD-DELMON. *Role of semantic vs spatial congruency in a bimodal go/no-go task*, in "International Multisensory Research Forum", July 2007, http://www-sop.inria.fr/reves/Basilic/2007/SBV07.

[24] N. TSINGOS, C. DACHSBACHER, S. LEFEBVRE, M. DELLEPIANE. *Extending Geometrical Acoustics to Highly Detailed Architectural Environments*, in "19th Intl. Congress on Acoustics", sep 2007, http://www-sop.inria.fr/reves/Basilic/2007/TDLD07b.

[25] N. TSINGOS, C. DACHSBACHER, S. LEFEBVRE, M. DELLEPIANE. *Instant Sound Scattering*, in "Rendering Techniques (Proceedings of the Eurographics Symposium on Rendering)", 2007, http://www-sop.inria.fr/reves/Basilic/2007/TDLD07.

[26] N. TSINGOS. *Perceptually-based auralization*, in "19th Intl. Congress on Acoustics", sep 2007, http://www-sop.inria.fr/reves/Basilic/2007/Tsi07.

References in notes

[27] O. DEUSSEN, C. COLDITZ, M. STAMMINGER, G. DRETTAKIS. *Interactive visualization of complex plant ecosystems*, in "Proceedings of the IEEE Visualization Conference", IEEE, October 2002, http://www-sop.inria.fr/reves/publications/data/2002/DCSD02.

[28] J. DORSEY, H. K. PEDERSEN, P. HANRAHAN. *Flow and Changes in Appearance*, in "ACM Computer Graphics (SIGGRAPH'96 Proceedings)", Aout 1996, p. 411–420.

[29] F. DUGUET, G. DRETTAKIS. *Robust Epsilon Visibility*, in "ACM Transactions on Computer Graphics (Proceedings of ACM SIGGRAPH 2002)", July 2002, http://www-sop.inria.fr/reves/publications/data/2002/DD02.

[30] F. DURAND, G. DRETTAKIS, C. PUECH. *The 3D Visibility Complex, a new approach to the problems of accurate visibility*, in "Rendering Techniques'96 (7th Eurographics Workshop on Rendering)", Springer Verlag, June 1996, p. 245–257.

[31] F. DURAND, G. DRETTAKIS, C. PUECH. *The Visibility Skeleton: A Powerful and Efficient Multi-Purpose Global Visibility Tool*, in "ACM Computer Graphics (SIGGRAPH'97 Conference Proceedings)", Aout 1997.

[32] F. DURAND, G. DRETTAKIS, C. PUECH. *Fast and Accurate Hierarchical Radiosity Using Global Visibility*, in "ACM Transactions on Graphics", vol. 18, April 1999, p. 128–170.

[33] T. FUNKHOUSER, I. CARLBOM, G. ELKO, G. PINGALI, M. SONDHI, J. WEST. *A Beam Tracing Approach to Acoustic Modeling for Interactive Virtual Environments*, in "ACM Computer Graphics (SIGGRAPH'98 Proceedings)", July 1998.

[34] T. FUNKHOUSER, N. TSINGOS, I. CARLBOM, G. ELKO, M. SONDHI, J. WEST. *Modeling Sound Reflection and Diffraction in Architectural Environments with Beam Tracing*, in "Forum Acusticum", September 2002, http://www-sop.inria.fr/reves/publications/data/2002/FTCESW02.

[35] C. LOSCOS, G. DRETTAKIS, L. ROBERT. *Interactive Virtual Relighting of Real Scenes*, in "IEEE Transaction of Visualization and Computer Graphics", vol. 6, October 2000, p. 289–305.

[36] C. LOSCOS, M. FRASSON, G. DRETTAKIS, X. GRANIER, B. WALTER. *Interactive Virtual Relighting and Remodeling of Real Scenes*, in "Rendering Techniques'99 (10th Eurographics Workshop on Rendering)", Springer Verlag, June 1999, p. 329–340.

[37] G. MILLER. *Efficient Algorithms for Local and Global Accessibility Shading*, in "ACM Computer Graphics (SIGGRAPH'94 Proceedings)", July 1994, p. 319–326.

[38] N. TSINGOS, T. FUNKHOUSER, I. CARLBOM. *Modeling Acoustics in Virtual Environments Using the Uniform Theory of Diffraction*, in "ACM Computer Graphics (SIGGRAPH 2001 Proceedings)", July 2001.

[39] N. TSINGOS. *Artifact-free asynchronous geometry-based audio rendering*, in "ICASSP'2001", May 2001.

[40] B. WALTER, G. DRETTAKIS, S. PARKER. *Interactive Rendering using the Render Cache*, in "Rendering Techniques'99 (10th Eurographics Workshop on Rendering)", Springer Verlag, June 1999.

[41] Y. YU, P. E. DEBEVEC, J. MALIK, T. HAWKINS. *Inverse Global Illumination: Recovering Reflectance Models of Real Scenes from Photographs*, in "ACM Computer Graphics (SIGGRAPH'99 Proceedings)", 1999.