The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2005) *Short Presentations*
M. Mudge, N. Ryan, R. Scopigno (Editors)

Graph Based Pre-Rendered Approach to Virtual Reality for Cultural Heritage Applications

A.F. Abate, M. Nappi, S. Ricciardi

Dipartimento di Matematica ed Informatica, Università di Salerno, Italy

Abstract
The term Virtual Reality usually refers to the promise of a visually believable interactive experience in a 3D world, typically visualized through real-time rendering of polygonal textured models. In this paper we present a different approach to the fruition of rich virtual environments populated by highly realistic computer generated characters, featuring state of the art lighting, shading and animation techniques. The proposed framework does not rely on real-time rendering but on high resolution stereo pre-rendering instead. A graph, whose nodes represent pre-rendered audio-video clips available for a given context, is visited during navigation and interaction, allowing the user to enjoy a believable virtual experience. This methodology, though limiting in term of action freedom could be effectively used for VR applications in fields such as Edutainment and Cultural Heritage. We show the proposed framework applied to the virtual exploration of the digitally reconstructed archaeological site of Moregine (Pompei) guided by one of its last inhabitants.

Categories and Subject Descriptors (according to ACM CCS): J.2 [Computer Applications]: Archaeology

1. Introduction

Since the first steps moved in the '70s, tri-dimensional computer generated imaging (CGI) has dramatically evolved, reaching an incredibly high level of realism simulating any kind of objects, environments and characters, either real or fictitious. Technical achievements in modelling, rendering and animation, have been the premise for a growing number of applicative fields, including design/prototyping, cinema special effects, scientific and architectural visualization, videogames and Virtual Reality (VR). Thanks to the technologically related revolution operated by Graphical User Interface (GUI) based operative systems, the above mentioned applications were empowered by (and partly responsible for) continuous innovation in both hardware (graphics boards, local/high speed buses, vector processing capable CPUs, multi-pipelined GPUs, etc.) and software (3D engines, VRML, OpenGL and DirectX graphics libraries, etc.). Indeed, 3D graphics intrinsic need for high speed computing has been one of the main reasons (at least from a technical point of view) for the research and the development of many architectural enhancement which characterize most computers today, so no wonder if the toughest PC or workstation benchmarks are based on 3D intensive tasks like polygonal based rendering.

Even if this scenario could suggest that time has come for photo-realistic tri-dimensional VR applications on every PC or even just on dedicated hardware, this is not true, at least not yet. Surely most advanced videogame engines coupled with the latest CPU and GPU technology can produce stunning graphics, but if polygonal detail and image resolution exceed a certain amount, like in large screen visualization for example, the illusion vanishes.

Even highly expensive dedicated hardware, such as multi-GPUs based solutions suited for real-time flight or astronomical simulation could not address highly detailed complex characters and environments rendered at cinema-quality by the most sophisticated shading, lighting and animation techniques. So, though certainly the next generations of specialized hardware will bring even more processing power to VR applications, at this moment a trade-off in terms of visual quality and/or scene complexity is still necessary. This study presents a VR framework offering virtually unlimited visual quality at the cost of a constrained interactivity, and results best suited for edutainment and cultural heritage applications like virtual museums, for example. It is based on a pre-rendered approach which relies on scalable, not dedicated hardware technology. We applied this framework to the reconstruction of the archaeological site of Moregine, near Pompei (see Fig 1.), and of one its inhabitants whose rests were founded there. This paper is organized as follows. In section 2 main approaches to VR are presented. In section 3 the proposed framework is presented in detail and applied to the case study. In section 4 the proposed methodology and its implementation is briefly compared to other approaches. The paper concludes in section 5.

2. Real-time, Image Based and Pre-rendered VR

Virtual Reality paradigm is to provide user a visually believable experience in a virtual environment, interacting within it with objects and characters. In a classic geometry based approach to VR, the interaction requirement asks for a high frame rate real time rendering, while the visual realism requirement asks for high image resolution, fine geometry detail and sophisticated rendering algorithms.

Other typical VR options include stereoscopic and/or multi-screen visualization, which are even more demanding in terms of processing load. Though technology is rapidly changing the horizon with fast paced announcements of new and more powerful vector processing units or multi-pipelined architectures for parallel vertex and pixel processing, to achieve at the same time a high level of interaction and cinematographic visual quality is still not possible. The result is that, except for the entertainment domain (videogames), real-time based VR often does not live up to its promise of realism. With a completely different approach, Apple Computer Inc. released in the '90s the Quick Time Virtual Reality (QTVR) Application Program Interface as an extension of its QuickTime video compression technology [C96]. In this Image Based Rendering (IBR) method, tri-dimensional visualization of an environment, is obtained projecting a panoramic view of it onto a cylinder (or a sphere or even a cube). This environmental map can be produced through image processing of actual wide angle photos stitched together or even by properly rendered synthetic images. Limited navigation option such as constrained camera rotations and movie branching are supported so it is possible to build a virtual world and within some extents to explore it, but the main limitation is that the scene has to be static, meaning that moving objects or characters are not allowed unless they are rendered into the scene in real time from conventional geometry based models. Other more sophisticated IBR approaches have been developed since then to model and render environments and particularly architectural scenes [DBY98], even combining geometry based and image based approaches [DTM96]. They proved to be very suited to real-time navigation in visually rich not animated worlds. IBR techniques have been also proposed to improve realism of complex object (statues, monuments, even characters) when they do not need to be seen closely [PGC03], or just to reduce the geometry processing load in real-time engines. While these methods can effectively contribute to enrich real-time navigation in virtual environment, they can not yet address effectively animated environments with complex characters. With the "pre-rendered" approach to VR the aim is to realize a virtual experience through a vast collection of pre-rendered video clips accessed on-the-fly according to user interaction. This idea is not new, indeed in the past, different technologies were proposed for interactive video content fruition when real-time rendering was simply not an option. The Interactive Video Disk (IVD), presented in the first half of '80s was the first optical support to allow the storage of a big amount of visual data either as still images or as broadcast quality video. The IVD capability to randomly access a particular frame or video clip within a fraction of a second, was responsible for its usage as interactive video platform in a few entertainment applications as well as for training and educational purposes, implementing a simple interaction on digitised footage [L80]. Since one of the most stable IT trends in the last decade has been price reduction of mass storage coupled with a huge increment in both memory space and performances, we have now the opportunity to revisit the IVD concept of video interaction overcoming many of its original limits. Indeed, is now possible to memorize and to play hours of uncompressed high definition video content, even stereoscopic, accessing any point of a given clip randomly without any noticeable lag. So, for applications in which visual realism is key, while complete freedom of action is not a strict requirement, pre-rendering could represent an interesting alternative to real-time VR. This is the idea behind the system described in the following section.

Figure 1. *The excavations in Moregine. As this site is not anymore open to visitors, it represents a valid candidate to show the potential of virtual reconstruction and fruition.*

3. The proposed framework

The aim of the proposed framework is the interactive fruition of high quality virtual worlds. Our main aim is to keep visual quality to the highest possible level, making content creators free to exploit even the most realistic (and computing expensive) techniques for modelling, animation and rendering, without any constraints in terms of scene or character complexity. On the other hand, since a pre-rendering based approach limits interactivity, we want to reduce this disadvantage through effective system design using smart techniques derived from videogame development. The whole system is exposed in the following subsections 3.1 to 3.3.

3.1. System's architecture

The foundation of our proposal is the development of a pre-rendered virtual world in which user can navigate following pre-built paths and performing available actions through a dedicated context sensitive system interface. In Fig. 2 a schematic view of system's architecture is shown.

The entire virtual world, is a collection of video clips stored on a high performance server, including every scene, eventually viewed from any allowed angle, and every path from one scene to another. Such a virtual world can be visualized as a graph whose nodes represents scenes and whose oriented arches represent available paths which take user from scene to scene. Any node or arch is uniquely identified with a tag pointing to the start frame of the corresponding video-clip. There can be various types of scene-nodes and various type of path-arches as shown in Fig. 3. Solid circled nodes, represent scenes in which user's point of view is fixed, dash circled nodes represent scenes in which the camera can rotate along one axis, while for scene corresponding to dot circled nodes the camera can rotate along two axis. Every arch coming out from a scene-node equals to an action available to user at that precise point of the virtual experience. Mono-directional arches only allow to move from source node to destination node according to the direction of the arrow, bi-directional arches represent a single clip that can be played in both direction, whereas two parallel mono-directional arches represent a path which is reversible, but only through two different clips. The heart of this system is the navigation engine, implementing a real time video editing application which, based on the virtual world graph and on user's input plays the correct video clip for a given context. Thanks to the underlying hardware technology (detailed in the following subsection), previously rendered clips are seamlessly played in the correct sequence. In this way the navigation and interaction results in a pseudo real-time experience. For instance, since the hardware architecture is able to play a video sequence in both direction at variable speed, user can smoothly change his viewpoint (on a given axis) via interface simply playing forward or backward a specific circular seamless pre-rendered animation (if that option is available for that particular scene-node).

Particular scene-nodes may contain "loop" animations which can be played waiting for a user's input. Clearly world and scene design plays a crucial role in this architecture, but the availability of huge and cheap mass storage allows up to several hundreds high resolution clips to be saved on a single server, for a potentially rich and varied virtual experience. A further advantage of this architecture is that content updating can be easily achieved simply changing each clip's tags. Moreover the system could provide different levels of content according to various kind of users, like tourists, students, children, etc. This can be done adding new layers of nodes and arches specifically tagged and therefore accessible only by the right type of user.

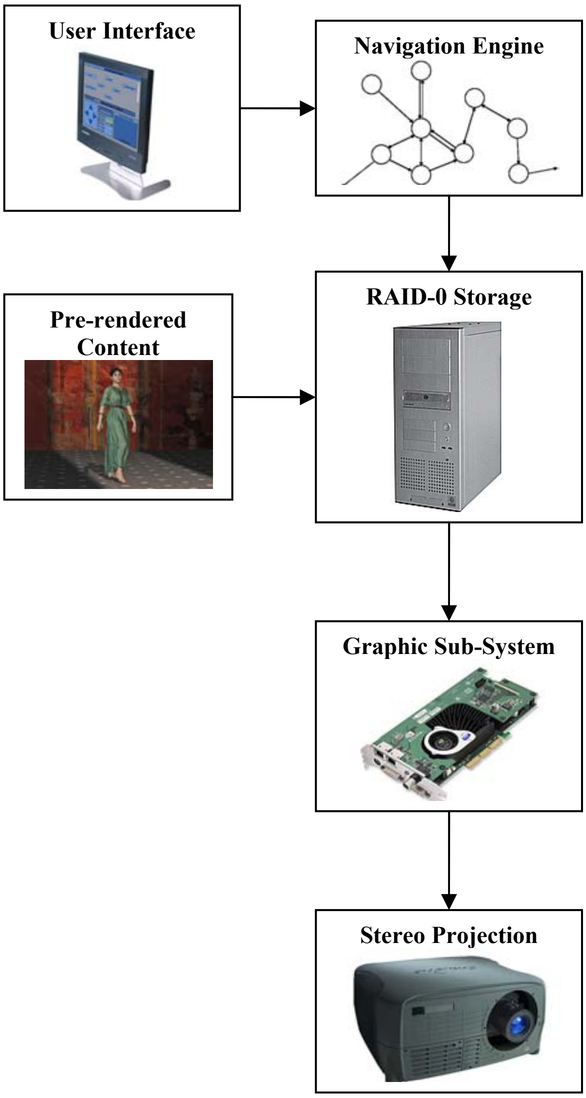

Figure 2. Schematic view of proposed system for interactive large-screen stereoscopic content fruition.

3.2. The hardware technology

The hardware equipments necessary for this architecture to work effectively, are tightly related to the requested image resolution and to the number of clips present in the Virtual-World-Graph (including all layers if any). Typical system configuration (for an output resolution up to 2K+ pixels, uncompressed frames) include a RAID-0 server equipped with an array of SATA or Ultra SCSI hard-disks, each one featuring high RPM and a 8-16 MBytes buffer for enhanced read and write operations. As RAID-0 operative modality implements a parallel access to all mounted drives this results in a single virtual disk whose capacity and transfer rate are the sum of the correspondent values of all physical drives. A total of 2-3 TBytes of storage space is not unusual allowing to play up to hours of high resolution uncompressed content with a sustained transfer rate of up to 300+ Mbytes/sec.. The navigation software run on the video server as it has a very low computing load since images are stored uncompressed, therefore it does not need special hardware or powerful processors. However, as high speed data exchange between storage sub-system and video frame-buffers is crucial for this architecture, 2 GBytes of RAM (used as content buffer) and a fast local bus (like PCI-Express) are required.

User interaction is accomplished through a LCD touch-screen, with a context sensitive graphical interface showing either actions or camera options available at a given stage of navigation. The user can select the desired action simply touching the correspondent button, while for each allowed camera movement directional arrows have to be touched. We are currently working on new and more intuitive concept of interface as briefly anticipated in section 5. The graphic subsystem is based on a workstation class graphic board (like the Quadro from Nvidia corp. or Fire from Ati corp.) which is not used for polygon processing but for its high pixel fill-rate, driving simultaneously up to two monitors or projectors in passive stereo, or one in active stereo through a quad buffered output.

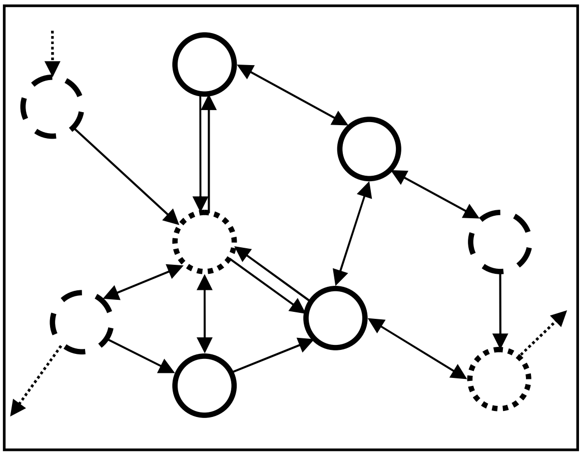

Figure 3. *A generic fragment of a Virtual-World-Graph.*

3.3. The case study

The proposed system has been used for the virtual reconstruction of the archaeological site of Muregine. This site is close to Pompei and so it was completely covered by ashes produced during the Vesuvio eruption on 79 a.C.. Excavations conducted in 1999 during expansion works on the local highway unearthed an entire beautifully decorated roman villa and its surroundings, together with a group of incomplete skeletons belonged to people who found death during the eruption. After a complete survey, wall paintings, objects and jewels were moved to the National Archaeological Museum of Naples and the site was newly filled with earth. Since at this moment nobody can visit the site anymore, VR represents the only way to see it again. One of the skulls found among the human remains in Moregine and belonging to a young woman have been the purpose of a face reconstruction research project [ANR*04]. We wanted to complete the reconstruction of this character, in order to use it as a guide in the virtual exploration of the site exploiting the potential of proposed system for sophisticated character animation. Thanks to the information provided by archaeologists, we realized a richly detailed digital replica of the site and its virtual inhabitant, exploiting the most advanced CG techniques like sub surface scattering for skin rendering, particle based hair, virtual garment design and simulation (see Fig. 4). As we did not need real-time performance, we were free to use even the most demanding lighting algorithms like radiosity and high dynamic range images (HDRI). The whole project has been optimised and rendered for active stereoscopic visualization.

4. Discussion

It is not difficult to imagine that in the next few years real-time CGI will deliver virtual experiences at near cinema quality with almost comparable content richness. After all, if we look at the most recent videogame titles, we could conclude it is just starting to happen. So it could seems a step back to propose a non real-time VR system in a time in which so many efforts are put in the research of real time solutions. But in the end VR is just the art of producing a believable illusion, no matter on what approach it is based on, so, why do not consider other available alternatives? Is reasonable to think that, if the illusion is broken, if real time CGI does not match the expectations, at least for some applications the interactivity alone could not be enough to capture user's interest (clearly this is not the case of videogames). But, on the other side, it is true as well that interactivity is what make VR an experience potentially more involving than others, like cinema for instance. So, until technology will not be ready to offer the best of both worlds, a trade off has to be accepted, and it could also be on the user's freedom of interaction, within the virtual experience. Probably, the term "freedom of interaction" itself is misleading even if it is referred to a genuine real-time simulation, because even in games each action the user can perform in a given moment (to open a door, to run, to shot, etc.) has been carefully planned from level designers, and usually can happen only under precise circumstances. What real-time VR really leave up to user (and even in this case there are some exceptions) is the freedom to move inside the environment looking everywhere he/she likes at almost any angle and zooming in and out. To this regard QTVR and IBR could produce photo-realistic environment visualization with a minimum hardware required, but they lack other form of interaction (i.e. to open a door is a very challenging task) and they are not suited for character animation.

Figure 4: *A frame from an animated clip showing the virtual guide for the reconstructed villa in Moregine.*

They could be used in a hybrid approach together to a real time system to partially overcome their limits while reducing the geometry necessary for the backgrounds. In Tab. 1 there is a resume of some of the characteristics featured in the three approaches to VR presented in section 2. As the comparison shows, real-time VR provides without doubt the highest level of interaction, but visual quality is directly related to rapidly changing specialized hardware, and cost can be prohibitive for large screen productions. Moreover the obsolescence of the required hardware and software is much faster resulting in a correspondent obsolescence of the content developed. With the proposed pre-rendered approach instead, content development is decoupled from hardware characteristics (except for image resolution and clips duration) and can even be integrated later simply adding new clips (new nodes and arches to the virtual world graph). Beside this, the cost of dedicated hardware (the RAID-0 server) is really low and is going to further reduce itself. As for a real-time VR application the level of optimisation required in every aspect of high quality content development is much greater then for pre-rendering, even the cost of the development can be higher. On the other side, the production of hours of pre-rendered content implies the use of render-farms or cluster of hundreds of pc, a number that is almost doubled in case of stereoscopic visualization. To this regard new IBR techniques have been recently developed to reduce considerably the extra time needed for ultra high resolution stereo rendering [SGH*01], making this option more attractive than ever. Concluding, pre-rendered VR can be effectively exploited in many contexts like virtual museums, reconstruction of historical characters, cultural heritage dissemination and fruition, which not necessarily need the advanced interaction and fast feedback of pure entertainment simulations, but could greatly benefit from unparalleled visual realism.

5. Conclusions and future enhancements

We presented a framework for pre-rendered VR, in which to navigate through the environment and to interact with its object and characters basically means to visit a graph. In this sense, nodes represent pre-rendered video clip of a scene in the virtual world, while oriented arches represent available paths, which seamlessly lead (through the appropriate animation) from one scene to another. Our aim is to not compromise visual quality, allowing any kind of high end modelling, animation, rendering and visualization techniques to be used for content production. Adaptivity to user's age or interests is another advantage offered by the proposed architecture, as the whole graph can be layered to satisfy different preferences.

The intrinsic limit of this approach, related to the pre-defined number of actions or viewpoints available in a given moment of the virtual experience, can be considered a well tolerable constraint in many applicative contexts and can be mitigated, adding new clips as, today, storage cost and space is not a bottleneck anymore.

Main Features	Approaches to VR		
	Real-Time	**QTVR/ IBR**	***Proposed***
Visual Quality	limited by hardware	high in static scenes	*virtually unlimited*
Interactivity	virtually unlimited	almost none	*limited by scene design / storage space*
Allowed Viewpoints	virtually unlimited	limited by source images	*limited by scene design / storage space*
Animated Objects or Characters	any, but performances are affected	no	*no limitation*
Large/Multi Screen Visualization	yes, but cost and performances are affected	yes	*yes*
Stereoscopic Visualization	yes, but performances are affected	yes	*yes*
Dedicated Hardware Cost	directly related to resolution / performances	none	*low, does not affect content quality*
Dedicated Hardware Obsolescence	fast, due to rapidly changing technology	No dedicated hardware required	*Slow*
Content Longevity	Directly Related to HW/SW	high	*high*

Table 1. *Resume of main features in approaches to VR*

The proposed framework produced interesting results on the case study, the stereoscopic virtual reconstruction of the archaeological site of Moregine and of one of its last inhabitants, now acting as a guide during the navigation, proving this approach to be really suited for (but not limited to) applications like virtual museums and historical reconstructions.

As interactivity is a fundamental aspect of VR we are working on an enhanced hybrid system interface, to provide a navigation/action interface featuring real-time visual feedback which operates through a cyber-glove. Rendering in real time the user's virtual hand(s) and the interface onto the pre-rendered scene could allow a much more intuitive and coherent way to select options and to interact with environment and characters compared to the separated interface screen adopted in this first implementation.

Acknowledgements

We would like to thank the Soprintendenza Archeologica of Pompei, for the precious help provided throughout the reconstruction project and Regione Campania which partly funded the research.

References

[ANR*04] A.F. Abate. M. Nappi, S. Ricciardi, G. Sabatino "3D Face Reconstruction from Skull Aimed to Archaeological Applications. The Site of Murecine: a Case Study", Proceeding of VAST2004: The 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, ACM/Eurographics, Brussels, Dicember, 2004, pp. 185-191.

[C95] S. E. Chen, QuickTime VR - An Image-Based Approach to Virtual Environment Navigation, Computer Graphics, Proceedings of SIGGRAPH 95, 1995, pp. 29-38.

[DBY98] P. Debevec, G. Borshukov, and Y. Yu. Efficient View-Dependent Image-Based Rendering with Projective Texture-Mapping. 9th Eurographics Workshop on Rendering, Vienna, 1998.

[DTM96] P. Debevec, C. J. Taylor, and J. Malik., Modeling and Rendering Architecture from Photographs: A Hybrid Geometry- and Image-Based Approach. In SIGGRAPH 96 Conference Proceedings, August 1996.

[L80] Lippman, A. Movie Maps: An Application of the Optical Videodisc to Computer Graphics. Computer Graphics(Proc. SIGGRAPH'80), pp. 32-43.

[PGC03] G. Papaioannou, A. Gaitatzes, D. Christopoulos, Enhancing Virtual Walkthroughs of Archaeological Sites, Proc. of ACM/Eurographics VAST 2003, Brighton, UK, 2003.

[SGH*01] H. S. Sawhney, Yanlin Guo, K. Hanna, R. Kumar, Sean Adkins, Samuel Zhou, Hybrid Stereo Camera: An IBR Approach for Synthesis of Very High Resolution Stereoscopic Image Sequences, Proceedings of SIGGRAPH'2001.