Virtual heritage, immersive environments and photorealism

Diego Gutierrez[1], Oscar Anson, Emilio J. Sobreviela, Francisco J. Seron

Grupo de Informática Gráfica Avanzada
Instituto de Investigación en Ingeniería de Aragón
Centro Politécnico Superior - Universidad de Zaragoza
C/ María de Luna, 1 - 50018 Zaragoza - Spain
Phone: +34 976 76 23 54 / Fax: +34 976 76 19 14
http://giga.cps.unizar.es
(1) diegog@unizar.es

1. Introduction

Simulation is the process that allows the study of a physical system, replacing it by another, more suitable for measure or observation. Some simulations work better using Virtual Reality (VR) techniques [1][2][3] where virtual worlds exist only in the memory of a computer. The user interacts with these virtual worlds by means of different types of devices (from 3D mouses to haptic devices), thus getting the sensation of being part of that world. VR is one of the technologies with the biggest potential in almost any area of application: medicine, architecture, scientific visualization and, of course, cultural heritage.

Within the context of the cultural heritage, VR allows us to go into past, present and even future worlds; we can walk inside long-gone buildings, interact with objects in ways that would be impossible in the real world, or study replicas of objects that are too fragile or hardly accessible in the real world, therefore helping preserve them untouched. We can test restoration techniques with a synthetic model, or explore different theories about its construction, while not damaging the original. The possibilities are limitless.

The goal of immersive environments is to provide the sensation of "being there" in the picture, by surrounding the viewers with images that belong to the simulated world. Those environments are closely related to VR applications, where the concepts of interactivity and real time come into play. In fact, those two concepts have taken over, and most of the immersive or VR applications focus on them. But there is another important factor that might need to be taken into account when trying to achieve immersion, and that is photorealism.

Given the power and speed of our current computers, having interaction, real time and photorealism in immersive environments is not feasible, specially in low-cost systems. Traditionally, photorealism has been discarded, in order to achieve the other two: interaction and real time. In other words, given the fact that we can not get all three at the same time, we usually lower the quality of the pictures until our computers are capable of rendering them in real time and account for some form of interaction. For certain applications, the end result is that we have replaced reality with our own version of reality.

2. Virtual Archaeology. Case study: Sinhaya

The term Virtual Archaeology describes the set of computer assisted techniques that allows the tridimensional visualization of realistic virtual replica of ancient objects and buildings, which real remains have disappeared, or else are in a state of preservation that makes its observation impossible or very difficult. In this context our group accomplished the digital reconstruction of a Muslim suburb from its remains found during an excavation.

In 2001, the Paseo de la Independencia, one of the main avenues of the city of Zaragoza (Spain), began being excavated to build a subterranean parking lot. Those excavations unexpectedly unearthed the remains of Sinhaya, a X-XII century Muslim suburb. After evaluating the different alternatives, and given the fact that keeping the remains visible to the public would have meant to strangle one of the city's main arteries, the decision was taken to abandon the parking lot project, while the Dirección General del Patrimonio Cultural arranged to have our group (GIGA) recreate the suburb in the digital realm. Using all the information recovered during the excavations, it would be shown to the public, in order to make citizens conscious about their own Heritage and to allow visitors to experience what was life like one thousand years ago. Our group then decided to render the images also in a different format, to be projected in stereo inside an immersive environment, and use that to test the role of photorealism (as opposed to interactivity and real time) in immersive environments. A more comprehensive explanation of the techniques employed can be found in [4].

To build our immersive environment, a low-cost CLS (CAVE-Like System) was chosen as output format [5][6]. Four to six people can fit in comfortably, although there is only one ideal point of view from which perspective and stereoscopy look perfect, since it is the point of view for which both have been calculated [7]. As we have said, the notion of real time was turned down in favor of high quality imagery. That means that the images have all been prerendered, which limits the interaction with the users (they can only jump from one movie to another, or pause) but in turns offers much better quality in the images themselves.

Our CLS is a theatre of 4x3x3 meters, made up of three rear projection screens for the front (3x3 m), right and left walls (4x3 m) and six LCD projectors, two projecting on each screen, with a 1024x768-pixel resolution. For stereoscopic vision we use polarized filters for each projector and their corresponding passive polarized glasses for the audience. To control the projection we use two mid-sized PC's; one is equipped with two MPEG-2 decoder cards, each one offering four channels of simultaneous MPEG-2 decoding, while the other one has six SVGA output channels to send static SVGA imagery to each projector. Figure 1 shows the CLS built. To keep costs low, it was decided not to have a floor or a ceiling screen. The room where the CLS is located has been painted in matt black to absorb any light filtering in or bounced off the screens. The system is built to be controlled from inside the CAVE, so a radio-based remote control is used.

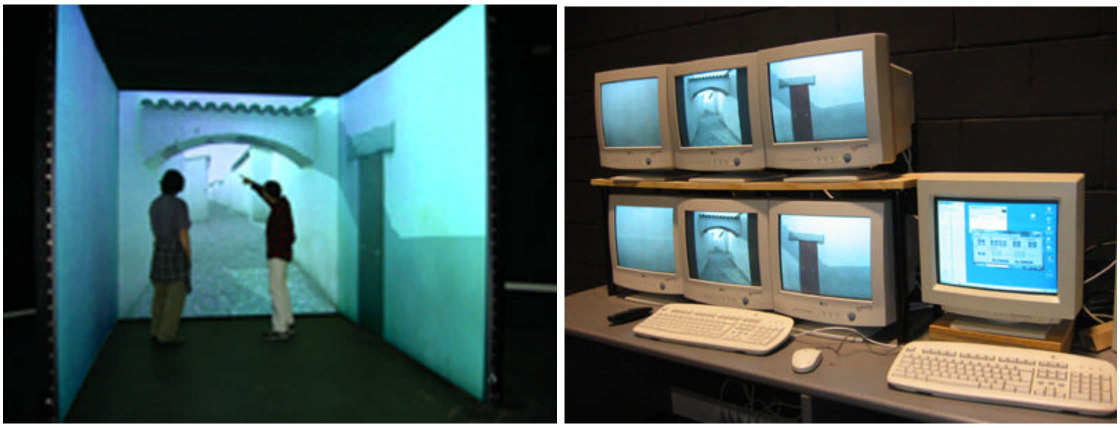

Fig 1. The CLS and its control room

Images and animations of the reconstructed Sinhaya can then be viewed in this patented CLS, designed and developed by GIGA [8]. The images are rear-projected on the three screens, thus accounting from most of the viewers' field of vision. Inside the CLS, the audience has therefore the feeling of being inside the suburb, with stereoscopy accentuating the illusion.

Almost five minutes of computer-generated imagery were delivered, whilst almost half of them were also rendered in CLS resolution to be used as our test on photorealism. The digital production follows a historically-accurate script, which takes the audience through the streets of the suburb in the morning into a pottery shop, a carpet shop, a wealthy man's house, then into a more modest house where the rest of the day passes by, and finally back to the streets again at night. Figure 2 shows several frames, both inside the CLS and as rendered images.

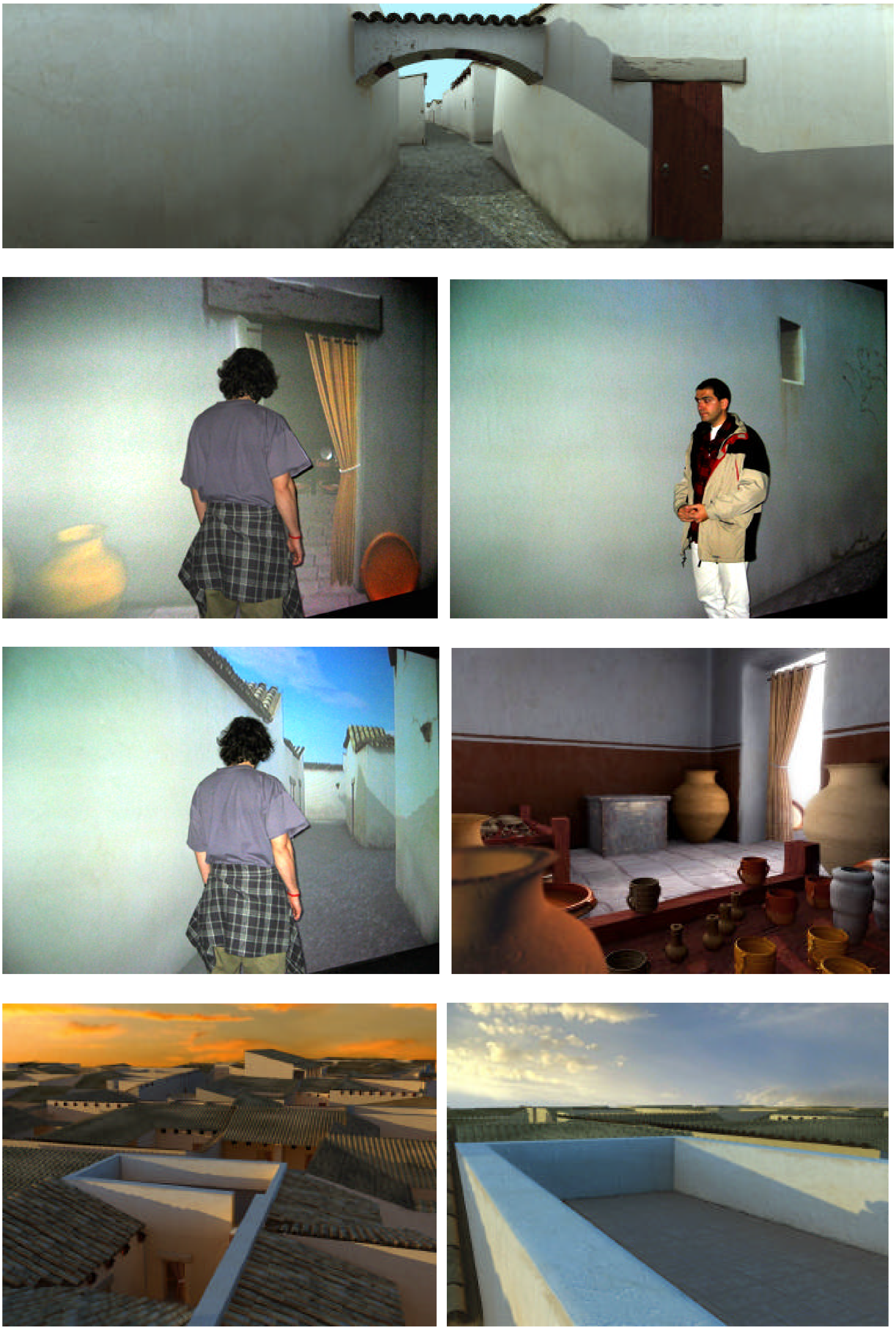

Fig. 2. Several frames of the reconstruction of the Muslim suburb of Sinhaya, in CLS format, inside the CLS and in VHS format

3. The dangers of Virtual Reality

We have to be careful avoiding spreading the notion that synthetic models can substitute real models [9]. No VR system, at least in a short-mid term, will be able to provide a simulation of the Great Wall capable of competing with the experience of actually travelling to the real Great Wall. It is really just a matter of bandwidth: the amount of information and stimuli that we process during the real trip (not only in terms of image quality, but also environmental noise, tiredness, cold or heat, wind … in short, all that makes real things *real*) can not be simulated by any computer system and much less in real time. A VR system has to therefore limit the amount of information (the bandwidth) to be transmitted, removing many of the sensations of the real world and keeping only those considered essential (usually, vision and sound). One more extreme example of bandwidth reduction in order to meet certain constraints is a chat session, as opposed to having a real conversation with a friend in a bar. The bandwidth has here been drastically reduced to a mere exchange of textual information in pseudo real time.

In the field of Virtual Heritage, specially when reconstructing buildings or cities as they were in ancient times, lowering the quality of images while favouring interactivity and real time might not always be the best option. We will probably obtain something close to the original, but with several important simplifications that the audience may not be able to make out. Imagine a digital reconstruction of a building or church to be shown in an immersive environment with interaction and real time. In order for our computers to be able to meet those two constraints, we will definitely have to simplify the model. This does not only imply the geometry of the model, but its textures and lighting algorithms as well, since there is nowadays no system in the world that can render a complex scene in a photorealistic way in real time, neither in terms of hardware nor software.

The result is therefore the substitution of reality for a *representation* of that reality, driven and limited by technological restrictions and personal interpretations and decisions. The audience, in general, will not be aware of those restrictions and decisions taken, and might then get the wrong impression. If this happens we will have failed in our initial goal of taking them to a long-gone past and showing them how it looked like. The power of technology and its capacity for transmitting knowledge has then turned against us. Historic rigour, to this effect, plays as important a part as technology itself, and if that means no real time, then we should consider dropping real time and going for a photorealistic look.

4. SEKER: Perception-based imagery

The goal of creating photorealistic synthetic imagery is to capture the visual appearance of the modelled scenes as exactly as possible. Physics-based rendering methods let us calculate the energy distribution in the scene with a great degree of accuracy. However, the exact calculation of this energy distribution does not guarantee that the visual appearance of the synthetic image matches that of the real scene, since the range of luminances in a real scene usually surpasses by several degrees of magnitude the dynamic range of an output device can handle. Pictures in a paper extend over a maximum contrast of 30:1; CRT monitors usually span no more than 100:1, and only a few high-quality prints can reach a 1.000:1 contrast ratio. In the real world, nevertheless, it is easy to find contrasts of over 100.000.000:1 or even more. Even though the projectors of our CLS have a dynamic range of 750:1, inside the CLS it goes down to the 50:1 range, owed to the filters, the distance from the projectors to the screens and the screens themselves. This 50:1 ratio is way below most of the real-world contrast ratios.

Tone reproduction provides a method of mapping (scaling) luminance values in the real world to a displayable range. The wide range of light in a real-world scene must be conveyed on a display with limited capabilities. This mapping is not obvious and several papers have been written addressing this issue [10][11]. In addition to compressing the range of luminance, tone reproduction is often used to mimic perceptual qualities of the human visual system (HVS). The goal is to generate an image that provokes the same physiological responses as when viewing the real scene in the real world. To do this, we must first obtain a model of this HVS, to include it in the rendering pipeline. Unfortunately, the studies on the matter have not been able to build a definitive and verifiable model yet, although different approaches have already been published [12] [13].

Yet another reason exists for which both human perception mechanisms and their application to tone reproduction are worth studying, and that is to save rendering times. Understanding how our brain is going to interpret the image, solutions can then be calculated with less precision from a physics point of view, but knowing that a more physically-accurate solution will not add anything to the image as perceived by a human observer, while also being more expensive in terms of rendering time. This is specially interesting in fields such as Virtual Reality or immersive systems such as the CLS, where the ultimate goal is to be able to produce photorealistic imagery in real time. We do not just want to see an image through an output device, however physically accurate it might be, but we want to create the feeling of actually being there instead, and that cannot be accomplished ignoring perceptual issues. Different perception-based metrics proposed by several authors can be found in [14] [15].

For instance, the visual perception of a scene varies considerably under different light conditions, mainly owed to the human visual limitations. Dark environments cause a loss of colour sensitivity and visual acuity, while bright sources produce veiling glare effects [16] [17] [18]. Advanced techniques of tone reproduction merged with models of human visual limitations should be employed if we want to simulate the perceptual sensation that real Muslim suburb would have produced in the past.

In this context, a tone reproduction application, as well as a model of the human visual system, was developed by the authors in a single environment called SEKER, and was tested in some of the animations of the suburb. Initial screenings inside the CLS proved that both a correct tone mapping and a simulation of several perception issues pay off in terms of increasing the sensation of realism inside the CLS.

The tone reproduction algorithm implemented in SEKER is based on iterative adjustments of the histogram of the image and spacial filtering processes. It uses more than one adaptation level for each image, since our visual system adapts over a visual area of 1º, not over the image as a whole. This area, also called foveal area, is the retinal zone where we focus all the visual attention, and the modification of the histogram is done from these adaptation groups. Such histogram is calculated from the logarithm of the average luminance in 1º areas over the image, using the foveal adaptation levels for each one. Only certain zones of the histogram are compressed, preserving contrast and details. Later, each real world luminance is efficiently mapped to display luminance to preserve visibility and contrast by means of the creation of a cummulative distribution function obtained from the modified histogram. The modification of the histogram tries to simulate the human sensitivity to contrast. The reader can refer to [18][19][20] for more specific details.

Some models of human visual limitations are also applied in the original algorithm, such as veiling glare, color sensitivity loss and visual acuity. Figure 3 shows a general scheme of the workflow of SEKER. It recieves High Dynamic Range (HDR) images [21], where each pixel has real world luminance values, and finally shows the image correctly tone mapped to the display luminance values.

Two criteria have been followed to make a reliable tone reproduction and, consequently, a better aproach to the desired photorealism. On the one hand, the reproduction of visibility; an object in the synthetic scene can be seen if and only if it can be seen in the real scene. Objects are correctly showed in both overexposed and underexposed regions. On the other hand, the vision of the image produces a subjective experience that matches the vision of the real scene. That is to say, what is showed in the display has to match the *memory* of the real scene. Bright, contrast and colour impression have then to be correctly reproduced.

Fig. 3. On top, the workflow of SEKER. Bottom, two tone mapped HDR images; left: only tone mapped, right: tone mapped and simulation of contrast sentivity, color loss, veiling glare and visual acuity loss.

5. Conclusions and Future Work

Photorealism is a valid option in certain VR applications, such as Virtual Heritage, even favored over real time or interaction. When historic rigour is a must, we have to be very careful not to replace reality with *our version* of reality. Initial screenings inside the CLS have proved that both a correct tone mapping and a simulation of several perception issues pay off in terms of increasing the sensation of realism inside an immersive system, in our case the CLS. The sense of immersion does not get diminished by the lack of interaction, if photorealistic imagery is offered instead. The audience will be ready to enjoy this form of passive VR, and concentrate on the imagery shown without missing more advanced forms of interaction. As it was shown during the conference, there are lots of different ways to visually depict Virtual Heritage, from hand-drawn 2D pictures to interactive 3D models to photorealistic renderings. All of them are valid, as long as we understand the visual language associated with each form of representation.

We have developed SEKER, an application that plugs into the rendering pipeline and reproduces several of the limitations of the human visual system while providing a good tone reproduction for different output devices, thus enhancing the photorealistic look of the imagery. On the other hand, the highly enthusiastic feedback from the users reinforces the idea that this system can be very useful in the promotion of cultural heritage and landscapes.

Of course there is still a lot of work to be done in the future. Our final goal is to be able to render photorealistic imagery with a model of the Human Visual System in real time, something that can not be done now. In that direction, we have implemented a Beowulf system [22] and are currently working on the parallelization of the rendering software, since raytracing scales pretty well. Also, there is still huge work to be done in avatar technology, to be able to populate the virtual reconstructions with photorealistic digital people that react and guide the audience, thus getting away from the dead-city look of most digital reconstructions.

6. Acknowledgements

This research was partially done under the sponsorship of the Comisión Interministerial de Ciencia y Tecnología (CICYT), project TIC 2001-2392-C03-02

7. Bibliography

[1] HOLLANDS, R., *"The virtual reality home brewer's handbook"*. John Wiley & Sons, Inc.: New York, 1996.

[2] BURDEA, G., *"Force and touch feedback for virtual reality"*. John Wiley & Sons, Inc.: New York 1996.

[3] DEL PINO, L.M., *"Realidad Virtual"*. Editorial Paraninfo, 1995.

[4] GUTIERREZ, D., SERON, F.J., MAGALLON, J.A., SOBREVIELA, E.J., LATORRE, P., *"Archaeological and cultural heritage: Bringing life to an unearthed muslim suburb in an immersive environment"*. Congress in Virtual Reality Applications. 2003, April 3rd-4th. Vitoria.

[5] CRUZ-NEIRA, C., SANDIN, D.J., DEFANTI, T.A., KENYON, R., HART, J., *"The CAVE: audio visual experience automatic virtual environment"*. Communications of the ACM 1992; Vol. 35, Num. 6, 64-72.

[6] CRUZ-NEIRA, C., SANDIN, D.J., DEFANTI, T.A., *"Surround-screen projection-based Virtual Reality: the design and implementation of the CAVE"*. Computer Graphics Proceedings, Annual Conference Series, SIGGRAPH'93, pp. 135-142

[7] MCALLISTER, D., *"Stereo computer graphics and other true 3D technologies"*. Princeton University Press, 1993.

[8] GUTIERREZ, D., SERON, F.J., MAGALLON, J.A., SOBREVIELA, E.J., GUTIERREZ, J.A., *"CLS: A low cost visualization environment for the train industry"*. Virtual Concept 2002. Biarritz (France), October 2002, ISBN2-9514772-2-8, pp. 122-127

[9] SPENCER, H., *"Supercharging the cultural engine: advanced media at heritage and educational attractions"*. En DODSWORTH, C. (ed) *Digital illusion*. ACM Press, SIGGRAPH Series, New York, 1998, pp. 489-500.

[10] TUMBLIN, J., RUSHMEIER, H., *"Tone reproduction for realistic computer generated images"*. Technical Report GIT-GVU-91-13, Graphics, Visualization & Usavility Center, College of Computing, Georgia Institute of Technology, Atlanta, GA 30332-0280. pp. 230-257. July, 1991.

[11] TUMBLIN, J., RUSHMEIER, H., *"Tone reproduction for realistic images"*. IEEE Computer Graphics and Applications, 13(6), pp. 42-48, November, 1993.

[12] ADELSON, E.H., *"Perceptual organization and the judgement of brightness"*. Science. Vol. 262, pp. 2042-2044. Diciembre 1993

[13] FERWERDA, J.A., *“Elements of early visión for computer graphics”*. IEEE Computer Graphics and Applications, pp. 22-33, Septiembre-Octubre 2001

[14] MCNAMARA, A., *“Visual perception in realistic image synthesis”*. ComputerGraphics Forum. Vol. 20, No. 4, pp. 211-224. 2001

[15] RADEMACHER, P., LENGYEL, J., CUTRELL, E., WHITTED, T., *“Measuring the perception of visual realism in images”*. In Paul Rademacher’s Ph.D. Thesis. 2001

[16] WARD, G., RUSHMEIER, H., PIATKO, C., *“A visibility matching tone reproduction operator for high dynamic range scenes”*. IEEE Transactions on Visualization and Computer Graphics, 3 (4), pp. 291-306. Octubre-noviembre 1997

[17] SPENCER, G., SHIRLEY, P., ZIMMERMAN, K., GREENBERG, D., *“Physically-based glare effects for digital images”*. Computer Graphics proceedings, pp. 325-334, ACM-0-89791-701-4/95/008. 1995

[18] FERWERDA, J.A., PATTANAIK, S.N., SHIRLEY, P., GREENBERG, D.P., *“A model of visual adaptation for realistic image synthesis”*. Computer Graphics Proceedings, Annual Conference Series, pp. 249-258. Siggraph 1996

[19] WARD, G., RUSHMEIER, H., PIATKO, C., *“A visibility matching tone reproduction operator for high dynamic range scenes”*. IEEE Transactions on Visualization and Computer Graphics, 3(4), pp. 291-306. October-November, 1997

[20] WARD, G., *“A contrast-based scalefactor for luminance display”*. In P.S. Heckbert (Ed.). Graphics Gems IV, Boston, Academic Press Professional. 1994.

[21] DEBEVEC, P., MALIK, J., *“Recovering High Dynamic Range Radiance Maps from Photographs”*. Proceedings of SIGGRAPH’97, Computer Graphics Proceedings, Annual Conference Series, pp. 369-378 (August 1997, Los Angeles, California). Addison Wesley. Edited by Turner Whitted. ISBN 0-89791-896-7.

[22] STERLING, T., BELL, G., JANUSZ, KOWALIK, B., *“Beowulf Cluster Computing with Linux, Scientific and Engineering Computation”*. MIT Press, October 2001