# The influence of cross-modal interaction on perceived rendering quality thresholds

V.Hulusic
Warwick Digital Lab
University of Warwick
United Kingdom
V.Hulusic@warwick.ac.uk

M.Aranha
University of Bristol
United Kingdom
aranha@cs.bristol.ac.uk

A.Chalmers
Warwick Digital Lab
University of Warwick
United Kingdom
A.G.Chalmers@warwick.ac.uk

## ABSTRACT

High-fidelity rendering is computationally demanding and has only recently become achievable at interactive frame rates on high-performance desktop PCs. Research on visual perception has demonstrated that parts of the scene that are not in the focus of viewer's attention may be rendered at much lower quality without this quality difference being perceived. It has also been shown that cross-modal interaction between visual and auditory stimuli can have a significant influence on perception. This paper investigates the limitations of the human visual system and the impact cross-modal interactions has on perceivable rendering thresholds. We show that by exploiting cross-modal interaction, significant savings in rendering quality and hence computational requirements can be achieved, while maintaining the same overall perceptual high quality of the resultant image.

**Keywords:** High-fidelity rendering, cross-modal, perceptual threshold.

## 1 INTRODUCTION

High-fidelity rendering is a process of computing accurate physically-based images using rendering algorithms based on real-world physical interaction. By taking into account complexity of the rendering process, e.g. simulation of photons' propagation in ray tracing algorithms, it is easy to infer that this process is very computationally demanding. Despite the improvements in performance of rendering related hardware and rendering algorithms, it is still not possible to render complex high-fidelity scenes at interactive rates.

Research on human vision and visual perception have demonstrated that the human visual system (HVS) has constraints which influence the way we perceive our environment [IKN98]. As a result of these limitations, it is common that high-fidelity rendered scenes have a greater level of details than it is possible to perceive. In order to quantify this so that it can be exploited within computer graphics, two phenomena were investigated: visual attention and inattentional blindness. It has been shown that by using these findings, the rendering process can been optimised to reduce the computational costs without perceivable degradation [CCW03] [YPG01].

Another factor when perceiving an environment is the affect of cross-modal interactions between visual and auditory stimuli [Mas06]. It has been demonstrated that this can be taken into consideration for improving the performance of rendering algorithms. Using sound it is possible to render animations at lower frame rates in conjunction with selective rendering techniques without decreasing the perceptible visual quality [MDCT05b] [MDCT05a].

In this paper we investigate the affect of cross-modal interaction on the perceived rendering threshold for high fidelity graphics involving an user study. In the following section, we shall present a short review of related work on perceptually-based rendering and cross-modal interaction. Section 3 talks about the stimuli and the experimental procedure and setup used in the user study. In Section 4 we analyse the results obtained from the experiment, comparing them with the results gained using Visual Difference Predictor (VDP) image comparison. Finally, we present our conclusions and suggest some ideas for future work.

## 2 RELATED WORK

Many researchers in the field of computer graphics have started to investigate the characteristics of HVS and the application of such features when developing rendering algorithms. Until now, the focus of this research has been visual attention, saliency and the application of these constraints for achieving the high-fidelity rendering in real time.

### 2.1 Visual Perception and Perceptually-Based Rendering

The way in which humans observe environments depends on mechanisms in the eye as well as visual pro-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Copyright UNION Agency - Science Press, Plzen, Czech Republic.

cessing in the brain. An important characteristic of the human eye is the angular sensitivity. Only objects located in the centre of the gaze can be perceived in full details, while toward the periphery the ability to perceive detail decreases. This phenomena, labelled as an “internal spotlight”, was introduced by James [Jam90] and further commented by Humphreys et al. [HB89]. This idea was further explored by Mack et al. [MR98], who formalised the concept of “Inattentional blindness”, in which items unrelated to a task fail to be perceived despite falling under the gaze of the viewer. This was demonstrated within the Computer Graphics field by Cater et al. [CCL02]. These results have been later used for developing perceptually-guided selective rendering systems, using different types of human point-of-gaze prediction [CDS06] [CCW03]. In such a system the scene is rendered so that the viewer is attending to at the highest quality while the remainder of the scene is rendered in a much lower quality, and thus at a substantially reduced computational cost, without the viewer being aware of this quality difference. Additional related work can be found in [FPSG97] [MTAS01] [Mys02] [RPG99] [SFWG04].

Related to this is the work of Itti et al. [IKN98], who developed the model of saliency-based visual attention which uses the saliency map, a two-dimensional map based on colours, intensity and orientations to predict the likelihood of regions drawing attention of the viewer. This can be used in order to decrease rendering computational costs. Similar approaches are used by Yee et al. using error tolerance map (Aleph map) in dynamic environments [YPG01]. Sundstedt et al. combined saliency and task map, creating a new selective guidance system called "Importance map" [SDL$^+$05]. An interesting approach using a “snapshot” image for generating a saliency map in real time was presented in [LDC05]. Luebke et al. [LH01] also presented a novel approach for reducing the model complexity using perceptual criteria.

Lately, some research has investigated perceived rendering thresholds for still images and animations, attempting to determine the limit at which further increase in rendering quality fails to become noticeable. Sundstedt et al. [SDC05] investigated perceived aliasing thresholds by altering the number of rays shot per pixel (rpp) on static scenes and animations. A similar study was conducted using small screen devices [ADCH06]. In both studies it has been demonstrated that it is possible to decrease the rendering quality and computational costs without a perceivable difference to an observer.

## 2.2 Cross-Modal Interaction

Despite the significant research into the limitation of the HVS within Computer Graphics, little work has been undertaken into the strong cross-modal interactions between visual and auditory stimuli. Research by Mastoropoulou et al. [MDCT05b] [MDCT05a] [Mas06] demonstrated that sound can influence the visual perception using knowledge from the psychology field in combination with selective rendering algorithms. This research inferred that by exploiting the presence of sound it is possible to render animations with lower frame rates without perceived difference in animation smoothness. It was also proven that observers are unable to perceive reductions in quality of animated sequences when selective rendering was used to render sound emitting objects (SEO) in higher quality than the rest of the scene.

Other research by Storms [L.98] and Winkler et al. [SC05] showed that using the high-quality audio stimuli increases the perceptual quality of video observed and that the quality of both of them contribute in the perceived quality, respectively.

There is currently even an EU project, CROSSMOD [CRO07] investigating the perception of audio and visual cross-modal interaction for real time rasterised graphics.

Although there is a growing body of work on the cross-modal effect of graphics and audio, no one, to the best of our knowledge, has investigated how the presence of audio may affect the perception of quality in a ray traced image.

## 3 EXPERIMENT

In order to quantify the perceived rendering threshold of high-fidelity images when sound was present we conducted a psychophysical study as well as analysing the perceptual quality of the results using the Visual Difference Predictor (VDP) [Dal93]. In our study we used the independent samples design. Our dependent variable was the perceived quality of the rendered images with the actual image quality and the audio background as the independent variables.

For the auditory stimuli three conditions were considered: no sound, related sound and noise. Every sound group consisted of 28 pairs of images, i.e. 7 images for each of 4 scenes. Every pair contained two images shown one after another in a five slide sequence, see Figure 1. All pairs were ordered randomly in order to avoid bias and reduce the affect of the human short-term memory (STM).

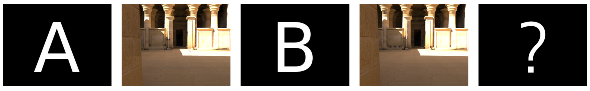

Figure 1: Example of the slide sequence from the experiment

Peterson et al., in their psychological studies [PP59] deduced that in presence of distractors, humans have problems with remembering even three elements for

Figure 2: Scenes used for the experiment (from left to right): Checkerboard, Corridor, Kalabsha, Library

more than eighteen seconds. Thus every picture in each pair was presented for 5 seconds, so that the whole sequence (Figure 1) lasted for about sixteen seconds.

### 3.1 Stimuli

For carrying out the research we rendered several 3D scenes which were used as visual stimuli and created audio samples for the accompanying sounds. To allow our work to be compared to previous work, we chose a sub group of scenes used by Sundstedt et al. [SDC05] and Aranha et al. [ADCH06], see Figure 2. Three scenes represented realistic environments and the fourth, the checkerboard scene, was used as a control scene because of its high spatial frequency.

For rendering we used the modified version of the *Radiance* `rpict` renderer [War94] developed by Debattista [Deb06]. This renderer was developed for selective, progressive and time-constrained rendering. The sampling algorithm is based on the hierarchical low-discrepancy (0,2) sampling sequence [KK02] composed of the Sobol and van der Corput sequences. Image reconstruction is performed using a Gaussian filter. This renderer is an improved version of the one used in the similar experiments used in [ADCH06, SDC05]. All scenes were rendered with 1, 4, 9, 16, 25, 36 and 49 rays per pixel (rpp), at a resolution of 1024 × 768 pixels. They were later converted to tif format using Radiance's `ra_tiff` command.

Using knowledge from previous studies by Mastoropolouou [Mas06], we decided to use related sound and noise, but not music. Sounds chosen for the experiment are presented in Table 1.

| Checker board | Corridor | Kalabsha | Library |
|---|---|---|---|
| Footsteps | Background chatting | Sounds of nature | Office noises |

Table 1: Related sounds used for the experiment

### 3.2 Visual Difference Predictor

In order to predict the probability of perceiving difference between the pair of images observed by a human, Daly developed the Visual Difference Predictor [Dal93]. This algorithm uses two images as input - the mask and a target image and compares them, producing a map of perceivable differences and values of detection probabilities (Figure 3). The VDP has some shortcomings: it can be used only for low dynamic range (LDR) images and it predicts the global level of adaptation to luminance. With significant progress in HDR applications and its general usage, Mantiuk et al. [MMS04] extended the original VDP algorithm to include comparison of HDR images taking into the account the local adaptation of the eye to every segment of the observed scene and the entire luminance spectrum visible to human eye.

Figure 3: VDP comparison. Top: mask image; Middle: target image; Bottom: difference map with probability of detection - green:0-50%; yellow: 50-75%; red:75-95%; pink:95-100%

Using the **HDR VDP**, we analysed an extensive range of 1024 × 748 images at varying rpp qualities. We chose a suitable gold-standard image beyond which there was no significant difference between images.

According to this method, it was found that at 1024 × 768, 49 rpp was an ideal gold-standard.

### 3.3 Hardware and rendering time

All images were rendered on a PC with a Intel Core2Duo E6650 CPU at 2.33GHz, 2GB of DDR2 PC6400 RAM memory and GeForce 8800GTS with 640MB graphic card. Rendering times are given in Table 2.

The experiment was conducted using Intel Pentium 4 computer working on 3.00GHz, Compaq 1825 19" monitor with 1280 × 1024 pixels resolution and Sony MDR-V300 sound-insulating Dynamic headphones.

| rpp | Checker board | Corridor | Kalabsha | Library |
|---|---|---|---|---|
| 1 | 13.12 | 673.99 | 98.29 | 130.57 |
| 4 | 51.98 | 1431.31 | 170.00 | 513.40 |
| 9 | 115.81 | 2471.16 | 288.62 | 1153.65 |
| 16 | 206.89 | 4110.91 | 409.61 | 2086.31 |
| 25 | 329.79 | 5556.64 | 581.24 | 3354.27 |
| 36 | 464.08 | 7807.02 | 796.40 | 4627.24 |
| 49 | 812.92 | 10327.32 | 1057.14 | 6277.47 |

Table 2: Rendering times for all scenes presented in seconds

### 3.4 Procedure

48 volunteers (33 male and 15 femal e) aged from 18 to 63, with an average of 25, participated in our user study. Most of them (about 85%) were undergraduate or postgraduate students at the University of Warwick. Other subjects were employees at the same University. All participants reported normal or corrected to normal vision with no hearing impairments. All use computers in everyday work, and reported average familiarity with computer graphics. The subjects were naive about the purpose of the experiment and participated in only one of the randomly selected group (no-sound, related sound, unrelated sound).

Subjects were asked to compare the quality of the images presented in each pair and choose the one they thought contained the higher rendering quality - two alternative forced choice (2AFC). The display of the stimuli was controlled using a program we wrote to provide fixed display time, synchronised audio and collection of the participants decisions. Images were displayed with a black background in a completely dark and noise-isolated room. The observers distance from the monitor was fixed at 60cm with subjects from all three groups asked to put on headphones. The experiment lasted about 9 minutes.

Prior to the experiment, participants were shown the demonstration with two image pairs which were not used in the study, the first of which accompanied with a related sound, and the second without. Images were displayed in pairs of the same scene and auditory stimulus with varying rendering levels. After each pair, a question mark would appear (Figure 1), in order to cue the participant to press either button 1 or 2 on the keyboard relating to which image they believe to contain the higher rendering quality. The keyboard was situated within arms-length distance from the observer on the desk in front, so there was no need to move the position of the body or the hand for pressing the buttons. After an answer was provided, the following pair of images were shown. All results were stored in a text file generated by the program.

After the experiment all subjects were asked to say on which features the differences were most obvious and approximately how many of the pairs they perceived any difference (i.e. they were not guessing).

## 4 RESULTS

Based on the research of Sundstedt et al. [SDC05], Aranha et al. [ADCH06] and Mastoropolouou [Mas06] we hypothesised that it would be possible to further reduce the number of rays shot per pixel without perceivable degradation when using audio stimuli as a distractor during scene observation.

In order to analyse the findings we used both *statistical analysis* of the psychophysical study as well as *comparison using the VDP*.

### 4.1 Statistical Analysis of Psychophysical Experiment

The results of our psychophysical study were analysed using *Pearson's Chi-square test* with *Yates' correction* [Yat34]. The Chi-square is a non-parametric test, commonly used to produce the statistical confidence of a hypothesis. This test allowed us to verify whether there was a statistical significant preference for one of the images in a pair. Our null hypothesis for each pair of images is that they should have equal preference. Since we assume no bias, it is expected that the gold-standard self test (49 v 49 rpp) should deliver equal preference. Our computed Chi-square values averaged over all scenes can be seen in Figure 4 with complete results presented in the Appendix.

The results of the Chi-square were analysed with a probability bound greater than 0.05 for significant results, i.e. a less than 5% chance in such an occurrence being observed.

According to this probability measure, a perceived degradation in rendering quality in comparison to our gold standard occurs in most scenes at less than 4 rays-per-pixel for the *no sound* and *related sound* groups. In the third group where we used *noise* as an audio stimulus, we can see that the subjects were unable to perceive any difference even at the 1 rpp comparison for the Corridor and Kalabsha scenes. For the other two scenes the

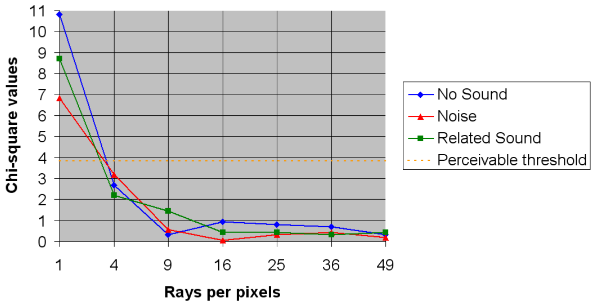

Figure 4: Comparison of chi-square thresholds averaged over all scenes

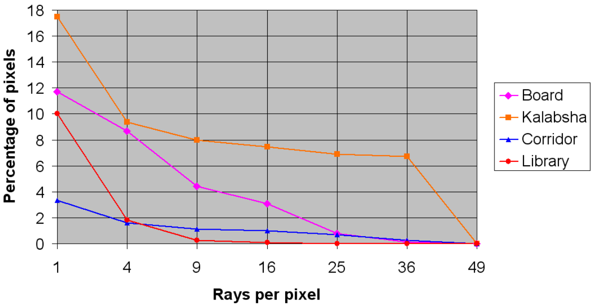

Figure 5: VDP comparison of perceivable differences for all scenes

perceived threshold was the same as in the *no sound* group.

The results were consistent across the majority of the scenes. As expected from the Checkerboard scene, the threshold was slightly greater than for the other scenes because of its high spatial frequency characteristics.

By comparing the results of the Chi-square across each of the three groups (no sound, related sound and noise), it can be seen that the *noise* threshold was lower than for the other conditions which confirmed our hypothesis. However, the *related sound* group contradicted our initial ideas and we actually experienced higher threshold level which may indicate that a related sound during the experiment caused subjects to look at the scene more closely.

It was found that in general the perceivable difference decreased monotonically with the increase in the number of rays-per-pixel.

## 4.2 Comparison using VDP

As discussed in Section 3.2, VDP can be used to highlight perceived differences as if viewed by the HVS. A constraint of VDP, however, is that it assumes significant viewing time. Our comparison between the psychophysical study and the VDP analysis allowed us to verify any difference which may have resulted from the finite viewing time.

In order to verify the results of our psychophysical study, we compared our gold-standard (49 rpp) with all other images of the same scene. The results of the VDP error measurements are included on Figure 5. These results were significantly different to what we experienced from our statistical analysis.

The results of our VDP comparison of the Kalabsha scene indicated significant variance between pairs of images, however, from our Chi-square results little statistical preference between images for this scene occurred. The differences between images in this scene are likely to be the result of texturing method of the model and the sampling strategy used by the renderer. This difference was picked up by VDP possibly due to HDR VDP's increased conservativeness, as described in the research by Ramanarayanan et al. [RFWB07]. For other scenes the perceptual threshold was significantly higher, in some cases more than ten times. We can also see from the Figure 5 that the difference was more likely to be perceived for the Checkerboard scene, as it was the case in our psychophysical experiment.

From post-experiment questioning of participants, approximately half of the subjects reported noticeable artefacts occurring on the edges of objects. This is verified by the VDP maps produced as seen in Figure 3. These regions are a consequence from the reduced sampling caused by tracing fewer rays in regions with high spatial variation.

## 5 CONCLUSION AND FUTURE WORK

Our experimental analysis indicates that cross-modal interaction has a significant affect on perceived thresholds. The presence of unrelated sounds, in our case noise, acts as a distractor as expected, producing a reduction in the thresholds. However, it was found that in the case of related sounds, a greater threshold was discovered which may indicate a heightened awareness of differences between rendering qualities.

We can conclude that for scenes with unrelated sounds, we can exploit such accompanying audio by ray tracing a lower number of rays without noticeable degradation resulting in reduced rendering times. Although our research indicates that related sounds may require an increase in rendering quality when compared to a no-sound scenario, we will investigate the manner in which this affects user's attention as a part of our future work. It may be possible that the focus is concentrated at specific regions, in which case it may still be possible to reduce rendering costs by using selective rendering techniques. As with any such statistical study, the validity of the result is dependent upon the sample size.

In the future we will further explore our findings by using a greater granularity of rpp levels around our discovered thresholds and a higher quality of the accompanying audio. We also intend to investigate the affect of cross-modal interaction on animations. The results of this research will form the basis of a cross-modally aware rendering framework.

## ACKNOWLEDGEMENTS

Thanks to Kurt Debattista for the use of his renderer and some general feedback on this work and Cheng-Hung Lo for his advice on experimental design and analysis. The use of the Checkboard, Corridor and Kalabsha was with the permission of the University of Bristol. The Library scene was created by Peter Apian-Bennewitz. We would also like to thank all the volunteers that participated in the experiment.

## REFERENCES

[ADCH06] Matt Aranha, Kurt Debattista, Alan Chalmers, and Steve Hill. Perceived rendering thresholds for high-fidelity graphics on small screen devices. In *Theory and Practice of Computer Graphics 2006*, pages 133–140. EG, July 2006.

[CCL02] Kirsten Cater, Alan Chalmers, and Patrick Ledda. Selective quality rendering by exploiting human inattentional blindness: Looking but not seeing. In *Symposium on Virtual Reality Software and Technology 2002*, pages 17–24. ACM, November 2002.

[CCW03] K. Cater, A. Chalmers, and G. Ward. Detail to attention: exploiting visual tasks for selective rendering. In *EGRW '03: Proceedings of the 14th Eurographics workshop on Rendering*, pages 270–280, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association.

[CDS06] Alan Chalmers, Kurt Debattista, and Luis Paulo Santos. Selective rendering: computing only what you see. In *GRAPHITE '06: Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia*, pages 9–18, New York, NY, USA, 2006. ACM Press.

[CRO07] CROSSMOD. Cross-modal perceptual interaction and rendering. World Wide Web electronic publication, 2007. `http://www-sop.inria.fr/reves/CrossmodPublic/`.

[Dal93] Scott Daly. *The visible differences predictor: an algorithm for the assessment of image fidelity*, pages 179–206. MIT Press, Cambridge, MA, USA, 1993.

[Deb06] Kurt Debattista. *Selective Rendering for High Fidelity Graphics*. PhD in Computer science, University of Bristol, 2006.

[FPSG97] James A. Ferwerda, Sumanta N. Pattanaik, Peter Shirley, and Donald P. Greenberg. A model of visual masking for computer graphics. *Computer Graphics*, 31(Annual Conference Series):143–152, 1997.

[HB89] Glyn W. Humphreys and Vicki Bruce. *Visual Cognition: Computational, Experimental and Neuropsychological Perspectives*. Lawrence Erlbaum Associates Ltd, East Sussex, BN3 2FA, UK, 1989.

[IKN98] Laurent Itti, Christof Koch, and Ernst Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(11):1254–1259, 1998.

[Jam90] William James. *The principles of psychology*. Holt, New York, 1890.

[KK02] T. Kollig and A. Keller. Efficient multidimensional sampling. *Computer Graphics Forum*, 21(3):557–563, September 2002.

[L.98] Storms R. L. *Auditory-Visual Cross-Modal Perception Phenomena*. PhD thesis, Naval Postgraduate School, 1998.

[LDC05] Peter Longhurst, Kurt Debattista, and Alan Chalmers. Snapshot: A rapid technique for driving a selective global illumination renderer. In *WSCG 2005*, February 2005.

[LH01] David P. Luebke and Benjamin Hallen. Perceptually-driven simplification for interactive rendering. In *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, pages 223–234, London, UK, 2001. Springer-Verlag.

[Mas06] Georgia Mastoropoulou. *The Effect of Audio on the Visual Perception of High-Fidelity Animated 3D Computer Graphics*. PhD in Computer science, University of Bristol, 2006.

[MDCT05a] Georgia Mastoropoulou, Kurt Debattista, Alan Chalmers, and Tom Troscianko. Auditory bias of visual attention for perceptually-guided selective rendering of animations. In *GRAPHITE '05: Proceedings of the 3rd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, pages 363–369, New York, NY, USA, 2005. ACM Press.

[MDCT05b] Georgia Mastoropoulou, Kurt Debattista, Alan Chalmers, and Tom Troscianko. The influence of sound effects on the perceived smoothness of rendered animations. In *APGV '05: Proceedings of the 2nd symposium on Applied perception in graphics and visualization*, pages 9–15, New York, NY, USA, 2005. ACM Press.

[MMS04] Rafal Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. Visible difference predictor for high dynamic range images. In Wil Thissen, Peter Wieringa, Maja Pantic, and Marcel Ludema, editors, *2004 IEEE International Conference on Systems, Man & Cybernetics*, volume 3, pages 2763–2769, Hague, The Netherlands, October 2004. IEEE.

[MR98] Arien Mack and Irvin Rock. *Inattentional Blindness*. The MIT Press, 1998.

[MTAS01] Karol Myszkowski, Takehiro Tawara, Hiroyuki Akamine, and Hans-Peter Seidel. Perception-guided global illumination solution for animation rendering. In Eugene Fiume, editor, *SIGGRAPH 2001, Computer Graphics Proceedings*, pages 221–230. ACM Press / ACM SIGGRAPH, 2001.

[Mys02] Karol Myszkowski. Perception-based global illumination, rendering, and animation techniques. In *SIGGRAPH 2002, Proceedings of the 18th spring conference on Computer graphics*, pages 13–24. ACM Press / ACM SIGGRAPH, 2002.

[PP59] L. R. Peterson and M. J. Peterson. Short-term memory retention of individual items. *Journal of Experimental Psychology*, 58:193–198, 1959.

[RFWB07] Ganesh Ramanarayanan, James Ferwerda, Bruce Walter, and Kavita Bala. Visual equivalence: towards a new standard for image fidelity. *ACM Trans. Graph.*, 26(3):76, 2007.

[RPG99] Mahesh Ramasubramanian, Sumanta N. Pattanaik, and Donald P. Greenberg. A perceptually based physical error metric for realistic image synthesis. In Alyn Rockwood, editor, *Siggraph 1999, Computer Graphics Proceedings*, pages 73–82, Los Angeles, 1999. Addison Wesley Longman.

[SC05] Winkler S. and Faller C. Audiovisual quality evaluation of low-bitrate video. *SPIE/IS&T Human Vision and Electronic Imaging*, 5666:139–148, 2005.

[SDC05] Veronica Sundstedt, Kurt Debattista, and Alan Chalmers. Perceived aliasing thresholds in high-fidelity rendering. In *APGV 2005 - Second Symposium on Applied Perception in Graphics and Visualization (poster)*. ACM, August 2005.

[SDL+05] V. Sundstedt, K. Debattista, P. Longhurst, A. Chalmers, and T. Troscianko. Visual attention for efficient high-fidelity graphics. In *SCCG '05: Proceedings of the 21st spring conference on Computer graphics*, pages 169–175, New York, NY, USA, 2005. ACM Press.

[SFWG04] William A. Stokes, James A. Ferwerda, Bruce Walter, and Donald P. Greenberg. Perceptual illumination components: a new approach to efficient, high quality global illumination rendering. *ACM Trans. Graph.*, 23(3):742–749, 2004.

[War94] Gregory J. Ward. The radiance lighting simulation and rendering system. In *SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 459–472, New York, NY, USA, 1994. ACM Press.

[Yat34] Frank Yates. Contingency table involving small numbers and the $x^2$ test. In *Journal of the Royal Statistical Society (Supplement) 1*, pages 217–235. Blackwell Publishing, 1934.

[YPG01] Hector Yee, Sumanita Pattanaik, and Donald P. Greenberg. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. *ACM Trans. Graph.*, 20(1):39–65, 2001.

## APPENDIX

| | Checkerboard | | Corridor | | Kalabsha | | Library | |
|---|---|---|---|---|---|---|---|---|
| No. of Rays | $x^2$ | $p$ | $x^2$ | $p$ | $x^2$ | $p$ | $x^2$ | $p$ |
| 1 | **7.562** | **0.0059** | **14.062** | **0.0001** | **7.562** | **0.0059** | **14.062** | **0.0001** |
| 4 | **7.562** | **0.0059** | 0.062 | 0.8033 | 0.062 | 0.8033 | 3.062 | 0.0801 |
| 9 | 0.062 | 0.8033 | 0.562 | 0.4534 | 0.062 | 0.8033 | 0.562 | 0.4534 |
| 16 | 0.062 | 0.8033 | 0.562 | 0.4534 | 0.062 | 0.8033 | 3.062 | 0.0801 |
| 25 | 0.562 | 0.4534 | 0.562 | 0.4534 | 0.562 | 0.4534 | 1.562 | 0.2113 |
| 36 | 1.562 | 0.2113 | 0.062 | 0.8033 | 0.562 | 0.4534 | 0.562 | 0.4534 |
| 49 | 0.062 | 0.8033 | 0.062 | 0.8033 | 0.562 | 0.4534 | 0.562 | 0.4534 |

Table 3: "No sound" group: Chi-Square Analysis (df=1; critical value 3.841 at 0.05 level of significance). Significant results in bold.

| | Checkerboard | | Corridor | | Kalabsha | | Library | |
|---|---|---|---|---|---|---|---|---|
| No. of Rays | $x^2$ | $p$ | $x^2$ | $p$ | $x^2$ | $p$ | $x^2$ | $p$ |
| 1 | **10.562** | **0.0011** | 3.062 | 0.0801 | 3.062 | 0.0801 | **10.562** | **0.0011** |
| 4 | **7.562** | **0.0059** | 0.562 | 0.4534 | 1.562 | 0.2113 | 3.062 | 0.0801 |
| 9 | 1.562 | 0.2113 | 0.562 | 0.4534 | 0.062 | 0.8033 | 0.062 | 0.8033 |
| 16 | 0.062 | 0.8033 | 0.062 | 0.8033 | 0.062 | 0.8033 | 0.062 | 0.8033 |
| 25 | 0.062 | 0.8033 | 0.562 | 0.4534 | 0.562 | 0.4534 | 0.062 | 0.8033 |
| 36 | 0.062 | 0.8033 | 0.062 | 0.8033 | 1.562 | 0.2113 | 0.062 | 0.8033 |
| 49 | 0.062 | 0.8033 | 0.562 | 0.4534 | 0.062 | 0.8033 | 0.062 | 0.8033 |

Table 4: "Noise" group: Chi-Square Analysis (df=1; critical value 3.841 at 0.05 level of significance). Significant results in bold. Different noise sounds: white, brown and pink noises were used in order to decrease the boredom and to avoid the bias as a consequence of the familiarization with the sound.

| | Checkerboard | | Corridor | | Kalabsha | | Library | |
|---|---|---|---|---|---|---|---|---|
| No. of Rays | $x^2$ | $p$ | $x^2$ | $p$ | $x^2$ | $p$ | $x^2$ | $p$ |
| 1 | **14.062** | **0.0001** | **10.562** | **0.0011** | **5.062** | **0.0244** | **5.062** | **0.0244** |
| 4 | 1.562 | 0.2133 | 0.562 | 0.4534 | 1.562 | 0.2133 | **5.062** | **0.0124** |
| 9 | **5.062** | **0.0244** | 0.062 | 0.8033 | 0.562 | 0.4534 | 0.062 | 0.8033 |
| 16 | 1.562 | 0.2133 | 0.062 | 0.8033 | 0.062 | 0.8033 | 0.062 | 0.8033 |
| 25 | 0.062 | 0.8033 | 0.062 | 0.8033 | 0.062 | 0.8033 | 1.562 | 0.2133 |
| 36 | 0.062 | 0.8033 | 0.562 | 0.4534 | 0.062 | 0.8033 | 0.562 | 0.4534 |
| 49 | 0.062 | 0.8033 | 1.562 | 0.2133 | 0.062 | 0.8033 | 0.062 | 0.8033 |

Table 5: "Related sound" group: Chi-Square Analysis (df=1; critical value 3.841 at 0.05 level of significance). Significant results in bold. Results show that the subjects were looking more closely, and were able to find more differences.