

Developing Effective Serious Games: The Effect of Background Sound on Visual Fidelity Perception with Varying Texture Resolution

David ROJAS^{a,f}, Bill KAPRALOS^{a,b,1}, Sayra CRISTANCHO^c, Karen COLLINS^d, Andrew HOGUE^a, Cristina CONATI^e, and Adam DUBROWSKI^{f,g}

^a*Faculty of Business and Information Technology,*

^b*Health Education Technology Research unit (HETRU),*

University of Ontario Institute of Technology, Oshawa, Canada.

^c*Department of Surgery and Centre for Education Research and Innovation, Schulich School of Medicine and Dentistry, Univ. of Western Ontario, London, Canada.*

^d*The Games Institute, University of Waterloo, Waterloo, Canada.*

^e*Dept. of Computer Science, University of British Columbia, Vancouver, Canada.*

^f*The Learning Institute, The Hospital for Sick Children, Toronto, Canada;*

^g*The Wilson Centre, University of Toronto, Toronto, Canada;*

Abstract: Despite the benefits associated with virtual learning environments and serious games, there are open, fundamental issues regarding simulation fidelity and multi-modal cue interaction and their effect on immersion, transfer of knowledge, and retention. Here we describe the results of a study that examined the effect of ambient (background) sound on the perception of visual fidelity (defined with respect to texture resolution). Results suggest that the perception of visual fidelity is dependent on ambient sound and more specifically, white noise can have detrimental effects on our perception of high quality visuals. The results of this study will guide future studies that will ultimately aid in developing an understanding of the role that fidelity, and multi-modal interactions play with respect to knowledge transfer and retention for users of virtual simulations and serious games.

Keywords: Visual fidelity, multi-modal cue interaction, audio-visual interaction, serious games, knowledge transfer.

Introduction

Recently there has been an increase in the use of immersive 3D virtual environments, particularly in the form of serious games. A serious game can be defined as interactive computer applications, with or without a significant hardware component, that i) has a challenging goal, ii) is fun to play and/or engaging, iii) incorporates some concept of scoring, and iv) imparts to the user a skill, knowledge, or attitude that can be applied to

¹Corresponding Author: Faculty of Business and Information Technology, University of Ontario Institute of Technology, 2000 Simcoe Street N., Oshawa, Ontario, Canada. L1H 7K4. bill.kapralos@uoit.ca.

the real world [2]. Serious games are able to captivate and engage players/learners (associated with academic achievement [8]), and allow users to experience situations that are difficult (even impossible) to achieve in reality due to factors such as cost, time, and safety concerns [9].

Despite the benefits of serious games, there are open, fundamental issues that must be addressed before their application becomes more widespread. There is a lack of research that investigates the technological conditions under which virtual learning can be maximized, as well as a lack of research that links virtual learning to proven educational theory and practice. More specifically, there are questions related to how fidelity and multi-modal cue interaction can affect immersion, knowledge transfer, and retention. Tashiro *et al.* [10] identified seven areas that require further research and improvements for the effective development of serious games, one of them being the impact of realism/fidelity on learning. In the context of a virtual learning environment or serious game, *fidelity* denotes the degree of similarity between the training situation and the operational situation which is simulated [4]. In other words, fidelity describes the extent to which the virtual environment emulates the real world [1]. *Transfer* can be defined as the application of knowledge, skills, and attitudes acquired during training to the environment in which they are normally used [6].

We have begun investigating the perception of visual fidelity of 3D rendered (virtual) scenes under various ambient auditory conditions. This is part of a larger-scale effort to examine virtual environment fidelity, multi-modal interactions, user-specific factors (e.g., personality, learning style, existing knowledge, level of attention, motivation etc.) and their effects on knowledge transfer and retention. The long-term goal of our work is to answer the questions i) “what effect do multi-modal interactions have on knowledge transfer and retention?”, and ii) “how much fidelity is actually needed to maximize transfer and retention?” The question have a number of implications when considering that any training device—be it a virtual or physical simulator—will never be able to completely replicate the real world, and the fact that in virtual worlds we have eliminated one sensory modality—smell, and reduced or restricted the haptic senses (touch and movement). Therefore, complete (perfect) multi-sensory fidelity appears to be impossible to achieve, at least with our current technology. In addition, striving to reach higher levels of fidelity can also lead to increased development costs, and it remains unclear if such fidelity is even needed for either enjoyment or knowledge transfer. Working towards our long term goal, we have begun conducting a number of studies that examine the effect of ambient (background) sounds on our perception of visual fidelity.

With respect to a 3D rendered (virtual) scene, visual fidelity can be defined in many ways, including polygon count and texture resolution. Our previous work examined the perception of visual fidelity with respect to polygon count under four auditory conditions: i) no sound at all (silence), ii) white noise, iii) classical music (Mozart), and iv) heavy metal music (Megadeth), and found that the perception of fidelity increased in the presence of classical music, particularly when considering images corresponding to higher polygon count [7]. Here, we build upon our previous work by examining the perception of visual fidelity under the same four auditory conditions, but we define greater fidelity with respect to higher texture resolution rather than higher polygon count. The results of this study will guide future studies that will ultimately aid in developing an understanding of virtual simulation and serious games fidelity, multi-modal interactions, user-specific factors and their effect on knowledge transfer and retention.

1. Methods & Materials

1.1. Participants

Participants consisted of volunteer students from the University of Ontario Institute of Technology (UOIT). A total of 18 volunteers (13 female, and 5 male) participated in the experiment (average age was 20 years old). No participants reported hearing or visual defects. The experiment abided by the UOIT Research Ethics Review process.

1.2. Visual Stimuli

The visual stimuli consisted of six images of a single surgeon's head against a black background (see Figure 1). The 3D model of the surgeon's head was comprised of 17,440 polygons but for each of the six images, as shown in Figure 1, fidelity was varied with respect to texture resolution only while all other parameters, including polygon count and image size, remained the same. The images were presented on an Acer Aspire laptop with a 15.6" screen size and a resolution of 1366×768 . The size of each image within the display was 800×630 . Each image was presented with its corresponding background sound until the participant made a choice. Participants were not able to interact with the image in any manner.

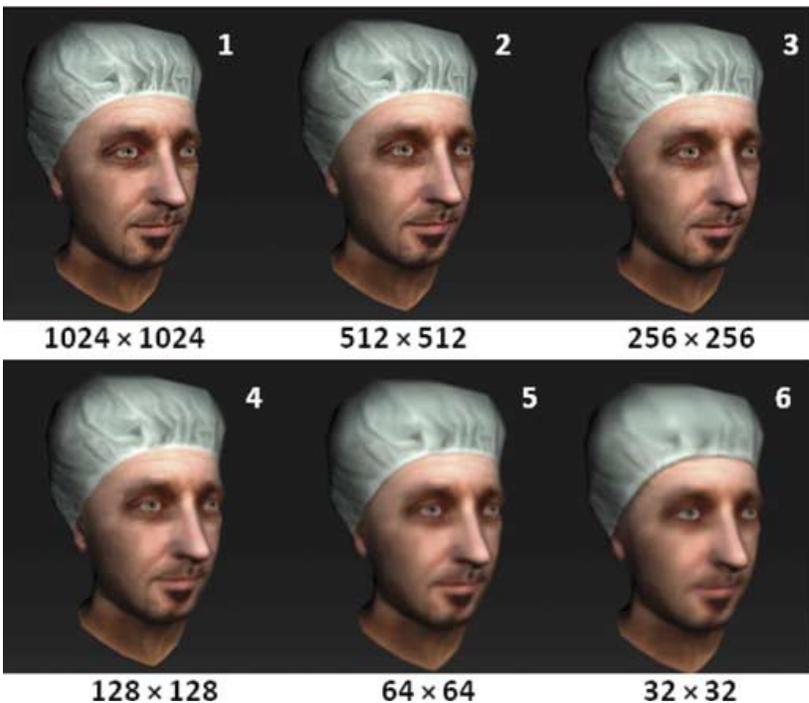


Figure 1. Surgeon's head textured with varying resolution textures.

1.3. Audio Stimuli

The auditory stimuli consisted of four background sound conditions: i) no sound at all, ii) white noise, iii) classical music (Mozart), and iv) heavy metal (Megadeth); these four auditory conditions are similar to the auditory conditions considered in the work described in [3,11] and in our own previous study [7]. The sound pressure level of the output sounds was 60dB (about the same level as typical conversation [5]) measured with a Radio Shack sound level meter (model 33-2055). All auditory stimuli were output with a pair of Sony MDR 110LP headphones.

1.4. Experimental Procedure

Prior to the start of the experiment, participants were presented with a brief questionnaire regarding their video game play habits and musical preferences. The experiment began immediately following the completion of the questionnaire. In each trial, participants were presented with one of the six images shown in Figure 1 in conjunction with one of the four background sound conditions described above. Their task was to rank the image with respect to their perceived visual quality on a scale from 1 (lowest perceived quality) to 7 (highest perceived quality). Each of the six images and each of the four background sound combinations (24 combinations in total) was repeated three times for a total of 72 trials.

2. Results

2.1. Questionnaire

The first question asked whether or not participants regularly played video games (in order to determine their regular use of computer-animated figures): 16 of the 18 participants (89%) played video games regularly. Of these participants, the average number of hours per week spent playing video games was 3.8 hours (minimum one hour and maximum 10 hours), six of them reported they primarily played console-based video games, seven of them primarily played PC-based video games, and three of them primarily played video games on their mobile phone. When asked to indicate their video game genre preference (from a list of given genres), six of the participants prefer “Strategy” games, five “Sports” games, four “Shooter” games, and one responded with “Other”. When asked to indicate their musical genre preference, five indicated “Metal”, three “Rock”, three “Country”, two “Pop”, two “House”, one “Arabic”, and one “Rap” (a list of musical genres was not provided to them; they had to indicate their own preference). Participant game and musical genre preferences did not appear to impact experimental results and therefore, will not be discussed further. The questionnaire then asked participants to rate several video game attributes according to their importance on a 7-point scale (1 being least important and 7 being most important). A summary of the average ratings for each of the attributes is provided in Table 1 and as shown participants believed that graphics and sound are the two most important attributes. Participant game and musical genre preferences did not appear to impact experimental results and therefore, will not be discussed further.

Table 1. Experimental results. Summary of video game attribute importance.

Video Game Attribute	Average “Importance”(/7.00)
Graphics/visuals	4.67
Sound	4.28
Interaction	2.89
Story/narrative	2.94
Level of challenge/difficulty	2.33

2.2. Experiment

A summary of the results is presented in Figure 2 where the x-axis represents texture resolution and the y-axis represents perceived quality (ranging from 1-7; error bars represent standard deviation). A visual inspection of Figure 2 indicates that there is generally a decreasing trend in perceived quality with decreasing texture resolution. A summary of the average, maximum, and minimum perceived quality across all texture resolutions for each of the background sound conditions is provided in Table 2.

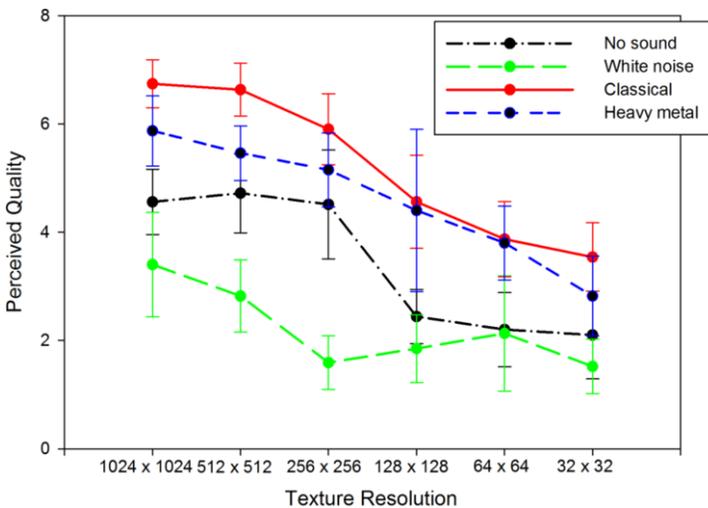


Figure 2. Experimental results. Texture resolution vs. perceived visual quality (with standard deviation).

Table 2. Experimental results. Average, maximum, and minimum perceived quality (across each of the six texture resolutions) for each of the background sound conditions.

Background Sound	Perceived Quality (Average ±Std. Dev.)	Perceived Quality(Max)	Perceived Quality(Min)
No sound	3.42 ±0.53	4.72	2.10
White noise	2.22 ±0.30	3.40	1.52
Classical	5.21 ±0.57	6.74	3.54
Metal	4.58 ±0.47	5.87	2.82

The analysis of variance (ANOVA) was selected as the statistical model with two factors: 6 images \times 4 background sounds. Main effects and interactions were further analyzed using Tukey HSD post-hoc comparisons. In addition to the main effects for image ($F=9.6$, $p<.001$) and sound ($F=21.2$, $p<.001$), there was a significant interaction between the sound and image ($F=5.8$, $p<.001$). Post-hoc tests revealed that the type of sound influenced the perceived quality of the images. Specifically, there was a similar scaling of the perceived quality of the image to the real quality of the image for the classical and metal (music) background sounds, which was not statistically different from the same as scaling with no background sound. That is, the presence of music did not affect the perception of image quality. However, white noise had an attenuating affect on the perception of the quality of the high-resolution images only but no effect on the perception of quality of the low-resolution images.

3. Discussion/Conclusions

In this paper we presented the results of an experiment that examined the perception of visual quality of 3D rendering/virtual model (defined with respect to texture resolution) under four auditory (background sound) conditions: i) no sound at all, ii) white noise, iii) classical music (Mozart), and iv) heavy metal (Megadeth). Results indicate that background sound consisting of white noise can have very specific and detrimental effects on the perception of the quality of high-resolution images (i.e., our perception of visual quality decreases in the presence of white noise when considering high fidelity visuals). In contrast, background sound consisting of music (classical or heavy metal) did not have any effect on our perception of visual quality. These results are supported by our previous work that also found that the perception of visual quality of a virtual model is dependent on ambient (background) sound [7]. However, in that work, visual fidelity was defined with respect to polygon count as opposed to texture resolution that was considered here and it was found that classical music increased perception of quality for the images corresponding to the highest polygon counts. Further evidence is also available from various “real-world” studies. For example, Woods *et al.* [11] discovered that background sound (noise) can have an effect on the perception of food gustatory properties (i.e., sugar level, salt level), food crunchiness and food liking. They found that background noise has three effects on food perception: i) food saltiness and sweetness was reduced in the presence of loud background noise, ii) food was “crunchier” in the presence of background noise, and iii) there was a correlation between the liking of the food and the liking of the noise. Conrad *et al.* [3] examined the effect of background sound (including classical music (Mozart), and heavy metal music to induce “auditory stress”) on laparoscopic surgery. They found that stressful music (e.g., heavy metal) had a negative impact on time until task completion but did not impact task accuracy. They also found that classical music had a variable effect on time until task completion but resulted in greater task accuracy amongst all participants (laparoscopic surgeons).

Collectively, the results presented here and in our previous study [7] suggest that the perception of visual quality of virtual models can be influenced by background sound. This is further supported by the real-world studies described above. This has implications for designers of virtual learning environments/serious games who should consider background sounds from the start of the design and development process particularly when high visual fidelity is sought.

This is a preliminary study, however, and there are many questions that must be considered before a greater understanding regarding the effect of fidelity and multi-modal interaction on knowledge transfer and retention is obtained. More specifically, with respect to the visuals, the sounds considered here lacked any context. What effect, if any, will contextual auditory cues have on our perception of visual fidelity? Moreover, it is unclear if specific attributes of the music were responsible for the results (such as the presence of lyrics, the frequency band, rhythm, and so on). Future work will consider examining the perception of visual fidelity with sounds that are contextualized to the visual scene presented. For example, with the visuals presented here (i.e., the surgeon's head), an additional auditory condition could be the background sound within an operating room during a surgical procedure. In addition to contextual auditory cues, future work will also consider the effect of user interaction on visual fidelity perception and, more specifically, actively involving the participant by having them perform a simple task within the virtual world under various auditory conditions (contextual and non-contextual). A simple measurement of performance could be defined and this will allow us to compare sounds that are unnatural to an interaction with a virtual object being acted upon with more natural sounds and will take us closer to answering our questions regarding fidelity and its effect on knowledge transfer and retention.

Acknowledgements

The financial support of the Canadian *Network of Centres of Excellence (NCE), Graphics, Animation, and New Media (GRAND)* initiative and the *Natural Sciences and Engineering Research Council of Canada (NSERC)* is gratefully acknowledged.

References

- [1] D. H. Andrews, L. A. Carroll, and H. H. Bell. The future of selected fidelity in training devices. *Educational Technology*, **35**(1996):32-36.
- [2] B. Bergeron. *Developing serious games*. Hingham, MA, USA: Thomson Delmar Learning. 2006.
- [3] C. Conrad, Y. Konuk, P. Werner, C. G. Cao, A. Warshaw, D. Rattner, D. B. Jones, D and Gee. The effect of define auditory conditions versus mental loading on the laparoscopic motor skill performance of experts. *Surgical Endoscopy*, **24**(2010):1347-1352.
- [4] R. T. Hays, and M. Singer. *Simulation fidelity in training system design*. New York: Springer. 1989.
- [5] C. Morfey. *The dictionary of acoustics*. Academic Press, San Diego, CA., USA. 2001.
- [6] P. M. Muchinsky (Ed.). *Psychology Applied to Work*. Stamford, CT: Wadsworth. 1999.
- [7] D. Rojas, B. Kapralos, S. Crsitancho, K. Collins, C. Conati, and A. Dubrowski. The effect of background sound on visual fidelity perception. *ACM Audio Mostly 2011*. Coimbra, Portugal, Sep. 7-9, 2011.
- [8] V. J. Shute, M. Ventura, M., Bauer, and D. Zapata-Rivera. Melding the power of serious games and embedded assessment to monitor and foster learning: Flow and grow. In *The Social Sciences of Serious Games: Theories and Applications*. U. Ritterfield, M. J. Cody, P. Vorderer (Eds.), pp. 295-321. 2009.
- [9] K. Squire, and H. Jenkins. Harnessing the power of games in education, *Insight*, **3**(2003):5-33.
- [10] J. Tashiro, and D. Dunlap. "The impact of realism on learning engagement in educational games," *ACM Future Play 2007*. Toronto, Canada, Nov. 15-17, 2007.
- [11] A. T. Woods, E. Poliakoff, D. M. Lloyd, J. Kuenzela, J. R. Hodsona, H. Gondaa, J. Batchelora, G. B. Dijksterhuis, and A. Thomas.. Effect of background noise on food perception. *Food Quality and Preference*, **22**(2011):42-47.