The Effect of Background Sound on Visual Fidelity Perception

David Rojas
University of Ontario Institute of Technology (UOIT)
2000 Simcoe Street North
Oshawa, ON, Canada.
david.rojas@uoit.ca

Bill Kapralos
UOIT
2000 Simcoe Street North
Oshawa, ON, Canada.
bill.kapralos@uoit.ca

Sayra Cristancho
University of Western Ontario
1151 Richmond Street,
London, ON, Canada.
sayra.cristancho@schulich.uwo.ca

Karen Collins
University of Waterloo
200 University Avenue West,
Waterloo, Ontario, Canada.
collinsk@uwaterloo.ca

Cristina Conati
University of British Columbia
2329 West Mall,
Vancouver, BC, Canada.
conati@gmail.com

Adam Dubrowski
SickKids Learning Institute
555 University Ave.,
Toronto, ON, Canada
adam.dubrowski@gmail.com

Copyright is held by the author/owner(s).
Audio Mostly 2011, September 7–9, 2011, Coimbra, Portugal.
ACM XXXXXXXXXXX.

Abstract
We have begun investigating multi-modal fidelity with respect to serious games and knowledge transfer relating to user-specific factors (e.g., personality, learning style, existing knowledge, level of attention, motivation etc). Here we describe an experiment that examines perceived visual fidelity (with respect to polygon count) under four auditory stimuli (ambient or background sound conditions: i) no sound, ii) white noise, iii) classical music, and iv) "heavy metal" music. Results indicate that perceived image quality is dependent on background sound particularly for the higher quality images where classical music leads to the greatest perceived quality. Although preliminary, results indicate that the type of background sound does influence perceived visual quality.

Keywords
Serious games, fidelity, audio/visual cue interaction, knowledge transfer, multi-modal.

ACM Classification Keywords
I.3.7. Computer Graphics: Three-Dimensional Graphics and Realism-*Virtual reality*. H.5.1 General: Multimedia Information Systems-*Audio input/output*.

General Terms
Human Factors. Experimentation.

Introduction

Serious games have been referred to as "games that do not have entertainment, enjoyment, or fun as their primary purpose" [7]. They may more formally be defined as an interactive computer application, with or without a significant hardware component, that i) has a challenging goal, ii) is fun to play and/or is engaging, iii) incorporates some concept of scoring, and iv) imparts to the user a skill, knowledge, or attitude that can be applied to the real world [2]. Serious games allow users to experience simulated situations that are difficult (even impossible) to achieve in reality due to factors such as cost, time, and safety concerns [3] and support the development of various skills including analytical and spatial, strategic, recollection, psychomotor skills, and visual selective attention [9]. Despite the benefits of serious games, there are open, fundamental issues related to immersion and knowledge transfer and more specifically, how fidelity, and levels of realism and multi-modal cue interaction can affect immersion, knowledge transfer, and retention [12]. These issues can be further compounded by user-specific factors such as personality, learning style, existing knowledge, level of attention, motivation etc.

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 [5]. In other words, fidelity describes the extent to which the virtual environment emulates the real world or how realistic the virtual world is [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 [11]. Our question is, how much fidelity is actually needed to maximize transfer? This question may have a number of implications, particularly when considering that any training device—be it a virtual or physical simulator—will never be able to faithfully completely replicate the real world with all its stimuli. In other words, complete (perfect) fidelity appears to be impossible to achieve, at least with our current technology. Furthermore, striving to reach full fidelity can also lead to increased development costs [10,12] and it remains unclear if such fidelity is needed for either enjoyment or knowledge transfer

In the real world, our senses are flooded with information. We are constantly receiving visual, auditory, vestibular, olfactory, and proprioceptic information and are able to integrate/process this information and respond immediately. Yet, replicating (what appears to be) a very simple action/response in a virtual world--particularly in real time--can be very difficult and computationally expensive.

We have begun investigating multi-modal fidelity with respect to serious games and knowledge transfer relating to user-specific factors (e.g., personality, learning style, existing knowledge, level of attention, motivation etc). The purpose of this work is to gauge perception of fidelity in multi-modal video games (that is, video games that include auditory and visual cues). At this stage, we are testing if audio does indeed influence fidelity perception. The information we obtain from this study will serve as the basis for more specific multi-modal interaction studies and ultimately will help us to develop more effective serious games and virtual simulations with respect to knowledge transfer and retention.

Deleted:

Here we describe an experiment that examines perceived visual fidelity in the presence of different auditory stimuli (ambient or background sound). Visual fidelity is defined here with respect to polygon count. Background sounds consisted of either i) no sound at all (silence), ii) white noise, iii) classical music, and iv) "heavy metal" music. Results indicate that the type of background sound does influence perceived visual quality. More specifically, our results show that image quality perception increased in the presence of classical music, particularly when considering higher quality (higher polygon count) visuals. Although preliminary, these results may have implications for designers of video games and serious games alike.

The Experiment

Participants

Participants consisted of unpaid volunteers and were either researchers or students from the University of Ontario Institute of Technology. A total of 18 (nine male and nine female) volunteers participated in the experiment (average age was 21). The experiment abided by the University of Ontario Institute of Technology Ethics Review process for experiments involving human participants

Visual Stimuli

The visual stimuli consisted of six images of a single surgeon's head against a white background (see Figure 1). As described in Table 1, the fidelity of each image varied with respect to polygon count only; everything else remained the same (texture resolution, etc.) but the number of polygons comprising the model of the surgeon's head was varied. When displayed, the image remained static (i.e., participant's were not able to interact with the image in any manner).

Figure 1. Visual stimuli. Six renderings of a surgeon's head, each varying with respect to polygon count.

Image/Model	Polygon Count
1	17,440
2	13,440
3	1,250
4	865
5	678
6	548

Table 1. Polygon count for each model in each of the six images.

Auditory 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 (similar to the work described in [4,14]). The sound pressure level of the output sounds was 60dB (about the same level as typical conversation [9]) measured with a Radio Shack sound level meter (model 33-2055).

Experimental Method

Prior to the start of the experiment, participants were presented with a brief questionnaire and were asked to answer several questions regarding their video game play habits and musical preferences. The experiment began immediately following the completion of the questionnaire. For 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 (total of 72 trials).

Results

Questionnaire

The first question asked whether participants regularly play video games or not: 14 of the 18 participants (78%) did play video games. For those participants, the average number of hours per week spent playing video games was 2.4 hours, six of them reported they primarily played console-based video games, six of them primarily played PC-based video games, and two of them primarily played video games on their mobile phone. 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 2 and as shown, graphics and sound received the highest scores (4.0 and 3.8 respectively) indicating participants believed they are the two most important attributes.

Video Game Attribute	Avg. "Importance" (/7.0)
Graphics/visuals	4.0
Sound	3.8
Interaction	1.7
Story/narrative	3.1
Level of challenge/difficulty	2.9

Table 2. Summary of video game attribute importance.

Experiment

A summary of the results for the first part of the experiment (where the images were presented individually), is presented in the plot of Figure 2 where the x-axis represents the six images (with varying polygon count) and the y-axis represented perceived quality (ranging from 1-7). A visual inspection of Figure 2 clearly indicates that there is a decreasing trend in perceived quality as polygon count decreases although for Image five (polygon count of 678) there appears to be a jump in perceived quality for all but the "white noise" sound condition. Furthermore, it is also evident that for the three higher fidelity images (the three images corresponding to polygon counts of 17,440, 13,440, and 1,250 respectively), participants perceived the image to be of higher quality in the

Deleted: and

Deleted: report playing

Deleted: that did report playing video games presence of background sound that consisted of classical music (Mozart).

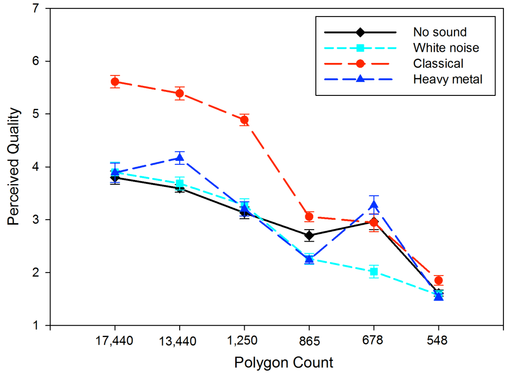

Figure 2. Results - image quality (polygon count) vs. perceived quality (images presented individually).

Discussion/Conclusions

Similar to the real-world, within a serious game and virtual environment, there are typically multiple sensory cues available. Sound, whether it is related to the task (i.e., task-specific) or not (i.e., ambient, or background sound) is typically present. As described above, in the real-world, it has been demonstrated that ambient/background sound does influence both our perception of the environment and the actions performed within the environment. For example, Woods et al. [13] investigated the effect of background sound (noise) on the perception of food gustatory properties (i.e., sugar level, salt level), food crunchiness and food liking. They found 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. [4] examined the effect of background sound (including classical music (Mozart), and dichaotic in the form of metal music to induce "auditory stress") on laparoscopic surgery. They found that dichaotic music 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).

Here we have presented the results of an experiment that examined perception of visual quality (with respect to polygon count) under four auditory (background sound) conditions: i) no sound at all, ii) white noise, iii) classical music, and iv) heavy metal. Similar to previous findings (e.g., see [4,13]), our results indicate that perception of visual quality can be influenced by background sound. More specifically, when considering the higher fidelity (higher polygon count) images, perception of visual fidelity increased in the presence of classical music.

Deleted: (Mozart)

Deleted: (Mozart)

Visual fidelity can be defined in many ways. Here we have chosen to represent fidelity with respect to polygon count but our future work will replicate this experiment but instead of defining visual quality/fidelity with respect to polygon count, we will define it with respect to texture resolution. Future work will also focus on the effect of ambient/background sound on actions and learning in more complex virtual environments that closely resemble a real-world scenario (e.g., an operating room).

Comment [k1]: What happened to my "future research" paragraph?

Acknowledgements

The financial support of the *Canadian Network of Centres of Excellence* (NCE), *Graphics, Animation, and New Media* (GRAND) initiative is gratefully acknowledged.

References

[1] D. H. Andrews, L. A. Carroll, and H. H. Bell (1996). The future of selected fidelity in training devices. *Educational Technology*, 35(6):32-36.

[2] B. Bergeron (2006). Developing serious games. Hingham, MA. USA: Thomson Delmar Learning.

[3] K. Corti (2006). Game-based learning; a serious business application, PIXELearning. Coventry, UK.

[4] C. Conrad, Y. Konuk, P. Werner, C.G. Cao, A. Warshaw, D. Rattner, D. B. Jones, D. Gee (2010). The effect of define auditory conditions versus mental loading on the laparoscopic motor skill performance of experts. *Surgical Endoscopy*, 24:1347–1352.

[5] R. T. Hays, and M. J. Singer, M. (1989). Simulation fidelity in training system design. New York: Springer-Verlag.

[6] V. Hulusić, A. Chalmers, K. Debattista, and V. Aggarwal (2011). Maintaining frame rate perception in interactive environments by exploiting audio-visual cross-modal interaction. *The Visual Computer: International Journal of Computer,* 27(1):*57-66.*

[7] D. Michael, and S. Chen (2006). Serious games: Games that educate, train and inform. Boston, MA. USA: Thomson Course Technology.

[8] A. Mitchell, and Savill-Smith (2004). The use of computer and video games for learning: A review of the literature. www.LSDA.org.uk. London, UK.

[9] C. Morfey (2001). The dictionary of acoustics. Academic Press, San Diego, CA., USA.

[10] M. Mori (1970). Uncanny valley. *Energy*. 7(4):33–35.

[11] Muchinsky, P. M. (Ed.) (1999). Psychology Applied to Work. Stamford, CT: Wadsworth.

[12] J. Tashiro, and D. Dunlap. The impact of realism on learning engagement in educational games. In *Proceedings of ACM Future Play 2007*. Toronto, Ontario, Canada, November, 2007.

[13] G. C. Visschedijk. The issue of fidelity: What is needed in 3D military serious games? Masters Thesis in Educational Science and Technology. *University of Twente, Faculty of Behavioural Sciences*, April 2010.

[14] A. T. Woods, E. Poliakoff, D. M. Lloyd, J. Kuenzela, R. Hodsona, H.Gondaa, J. Batchelora, G. B. Dijksterhuis, A. Thomas (2011). Effect of background noise on food perception. *Food Quality and Preference*, 22(1):42-47.