

Exploiting Audio-Visual Cross-Modal Interaction to Reduce Computational Requirements in Interactive Environments

Vedad Hulusić, Kurt Debattista, Vibhor Aggarwal and Alan Chalmers
International Digital Laboratory, WMG
University of Warwick, UK
Email: v.hulusic@warwick.ac.uk

Abstract—The quality of real-time computer graphics has progressed enormously in the last decade due to the rapid development in graphics hardware and its utilisation of new algorithms and techniques. The computer games industry, with its substantial software and hardware requirements, has been at the forefront in pushing these developments. Despite all the advances, there is still a demand for even more computational resources. For example, sound effects are an integral part of most computer games. This paper presents a method for reducing the amount of effort required to compute the computer graphics aspects of a game by exploiting movement related sound effects. We conducted a detailed psychophysical experiment investigating how camera movement speed and the sounds affect the perceived smoothness of an animation. The results show that walking (slow) animations were perceived as smoother than running (fast) animations. We also found that the addition of sound effects, such as footsteps, to a walking/running animation affects the animation smoothness perception. This entails that for certain conditions the number of frames that need to be rendered each second can be reduced saving valuable computation time. Our approach will enable the computed frame rate to be decreased, and thus the computational requirements to be lowered, without any perceivable visual loss of quality.

I. INTRODUCTION

High-fidelity graphics are required if serious and entertainment games are to realistically reproduce the same visual quality of the real world. In the case of serious games this is particularly important for (re-)creating environments that can be used for accurate training purposes in a number of fields, such as engineering, architecture, planning, defence, cultural heritage etc. Failure to provide the appropriate level of realism for training in virtual environments can result in the user subsequently adopting the incorrect response when confronted by the same event in a real situation.

Rendering realistic high-fidelity graphics in real-time is still unachievable on high-end desktop machines. Along with the visuals, audition is frequently being used in virtual environments to heighten the sense of realism. In addition, other senses such as olfaction are now also beginning to be explored in virtual environments [1], [2].

The expense of computing virtual environments, particularly those involving multiple senses, has led researchers to explore the interactions between different senses within the sphere of the Human Sensory System (HSS). Despite being extremely complex, the HSS is not perfect and it has certain limitations.

In particular, limited attentional resources ensure that the HSS cannot attend to all aspects of all the senses concurrently. Perception across modalities is major topic of research in psychology, for example [3], [4], [5], [6], [7], [8], [9]. Based on such work, one particular cross-modal effect, which has been successfully exploited in the field of computer graphics is that of the vision and audition [10]. Mastoropoulou et al. [11] demonstrated how audio could be used as a distractor to maintain high quality around a sound-emitting object while reducing the quality of the graphics in other areas in the context of offline rendering for animations. Mastoropoulou et al. [11] also demonstrated how sound effects could be used to reduce the computed frame rate of an animation without the participants perceiving any difference in the visual quality of the animation. Hulusic et al. [12] demonstrated similar results using artificially created beats within the context of animations.

While such methods have been successful, the sound effects used are usually unrelated to the actions occurring within the environment. In this paper, we attempt to demonstrate the efficacy of such cross-modal methods using an affect strongly associated with the movement being performed. We conducted a psychophysical experiment to determine whether sound effects arising from the footsteps of a participant affect the perception of the smoothness of an animation, under the conditions of running and walking. Our results show that this is indeed the case. We also studied how the effect of the speed of the camera, representing running and walking, affects the perception of smoothness of the animation. These results have implications for games, where the computation requirements for any scene may depend on the actions of the camera. For example in a first-person or third-person game, when running, more of the environment can come within view within a smaller time frame. This results in additional computation to render a single frame due to the need to render more geometry and textures. In such cases it may simply not be possible to achieve a required frame rate of 60 fps on existing hardware. With our method, this may not be necessary.

The rest of the paper is organised as follows. Section II gives an overview of the previous work done on this topic both in psychology and computer graphics. In Section III we explain the design, methodology, procedure and study question of the

psychophysical study. The results are presented in Section IV. Finally, in Section V we discuss and conclude the paper and provide some directions for future work.

II. RELATED WORK

The perception of individual senses has been studied for more than one hundred years [13]. A number of limitations have been identified with the Human Visual System (HVS), such as Inattentional Blindness [14], [15], and auditory system such as Continuity Illusion phenomenon [16] or auditory masking [17] also known as *the cocktail party effect* as well as factors which effect both sound and vision including angular sensitivity or internal spotlight [13], [18].

Cross-modal interactions, however have only been a common research topic in psychology in the last few decades. Although some findings, such as [19], [20], [3], [5], [8], [7] have been made, there are still many unanswered questions. Computer graphics researchers have also recently started considering these phenomena in order both to increase rendering performance and to achieve more accurate depictions of real world scenes in virtual environments.

A. Auditory-visual cross-modal interaction research in psychology

Although it has been shown that cross-modal interaction exists between different modalities such as smell and vision [21], touch and vision [22] in this paper we will focus only on the auditory-visual cross-modal interaction. It is important to outline that the interaction between audition and vision works in both directions: as audition influences vision, vision can also have effect on auditory perception. The foundations for this was presented in the research by Welch et al. [23] who introduced the *modality appropriateness hypothesis*. This hypothesis states that the modality that is more appropriate for a particular task will dominate the perception in that particular task. In other words, vision will dominate perception in the spatial domain because of its higher acuity, while in the temporal domain audition will be dominant modality.

The best example of the visual dominance over audition is the ventriloquism effect [19], [24], [25], [26]. This effect shows that humans, while watching TV or a puppet show, associate a sound source to a speaking person/puppets mouth even though it originates from the speaker/ventriloquist positioned at a different location. Another example is the McGurk effect [20] where a sound of /ba/ is perceived as /da/ when accompanied with lip movement corresponding to the pronunciation of /ga/.

Of more interest for this work is the effect audio has on visual perception. There are few known phenomena which we will briefly explain here. The first one is the auditory driving effect [3], [4], [27], which shows that, when presented simultaneously, sound *drives* vision in the temporal domain. Another phenomenon investigated by Shams et al. [5], [6] is the illusory flash induced by sound. This effect illustrates how an illusory flash can be induced by a sound beep, where we, for example, observe two flashes accompanied by three beeps.

Analogous to the ventriloquism effect, Morein-Zamir et al. [8] introduced the temporal ventriloquism. This effect shows that the time between two light flashes is perceived longer when a sound is presented before the first and after the second light, but shorter when there are two sound beeps between the lights. Some other examples can be found in [28], [7], [29], [30], [9].

There are alternative theories, such as the central and divided attentional resources theories [31], [32], [33], [34], [35], [36], [37] which look at how our attentional capacity can affect perception.

B. Auditory-visual cross-modal interaction research in computer graphics

Perceptual methods, such as perceptually-based rendering [38], [39], have over the past few years frequently been used within computer graphics. Similarly, perceptual factors were considered in some research on audio rendering [40], [41], [42]. However, this research focuses on a single modality and the factors influencing that modality.

Although the auditory-visual cross-modal interaction in computer graphics is becoming an increasingly popular research topic, this work is still very much in its infancy. One of the first to tackle this problem was Mastoropoulou in her PhD thesis [10]. The auditory influence on visual perception can be investigated from two perspectives: the spatial and the temporal. In the former, focus is on the spatial image quality while the latter focuses on temporal quality such as frame rate perception.

Mastoropoulou et al. [43] showed that selective rendering can be efficiently used so that only sound emitting objects are rendered in high quality and the remainder of the scene in much lower quality, without any significant difference in visual quality. Hulusic et al. [44] examined how related and unrelated audio influences visual perception and showed that unrelated sound can be used for increasing the perceptual quality of graphics.

Mastoropoulou et al. [45] investigated how music can influence perception of frame rate and the perceived duration of a video animation. Subsequently, they investigated how sound effects, e.g. a phone ringing or a thunder clap can be used as a distractor to vision [11]. This work showed that when accompanied by these scene-unrelated sound effects it is possible to decrease the frame rate of a video animation without perceivable difference in visual quality. Most recently, Hulusic et al. [12] focused particularly on the effect of beat rate, scene and familiarity on the perception of frame rate. They showed that in case of static scenes lower beat rates have a significant effect on perception of low frame rates.

III. EXPERIMENTS

In this study number of factors which could potentially influence the perception of the animation smoothness: frame rate, camera speed and sound effects were investigated. First we looked at the difference in smoothness perception of the slow - *walking* and fast - *running* animation. Our research hypothesis was that the speed of the camera movement will

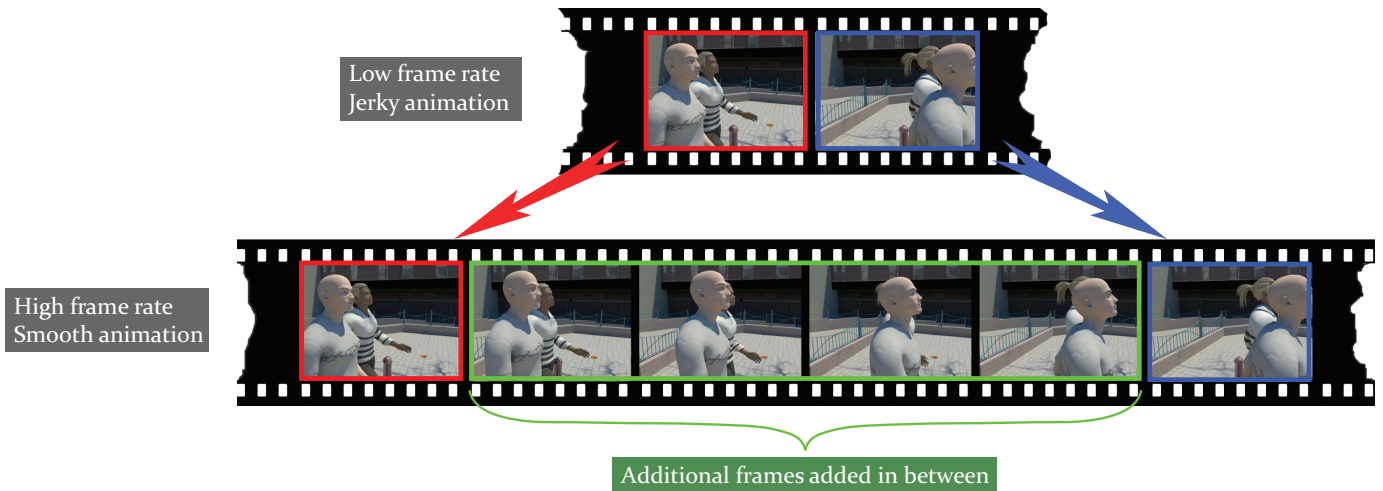


Fig. 1. Illustration of the difference between the low and the high frame rate.

affect the smoothness perception. Secondly, we investigated the perceived smoothness threshold for the animations accompanied by the audio effects. Our research hypothesis was that there will be difference between the preference of the discrepant frame rates and preference of the control condition.

A. Design

The experiment used a within-participant design. Three independent variables were used: camera movement, frame rate and auditory condition. There were two camera conditions: *walking* (slow) and *running* (fast). Four different frame rates were used: 10, 20, 30 and 60 frames per second (fps), in the following combinations: 10vs60, 20vs60, 30vs60 and 60vs60 fps were compared, see Figure 1. Lower frame rates were always presented for the running animations. Audio conditions were: *Audio* (foot steps) and *noAudio* (silent animation). The dependent variable was the perceived smoothness of the animations. This was measured using the Two Alternative Force Choice (2AFC) method in a complete randomised design. To control for fatigue and familiarity, 10 different animations of the same scene were used. In all test pairs, the presented animations were from different camera condition (walking and running). For Audio condition the foot steps were always synchronised with the visual stimulus.

B. Participants

In the experiment 36 people volunteered, 28 of whom were university students studying a variety of subjects, and the rest from university staff. The participants' age varied from 19 to 58 with an average age of 28. Out of 36 participants, 23 were male and 13 female. All of them had normal or corrected to normal vision. None of the participants reported any hearing impairments.

C. Apparatus

The experiment was conducted in a dark, quiet room with no surrounding distractors. Visual stimuli were presented on a

17 inch calibrated Philips 170B6 monitor at 1280×1024 pixels resolution at 60Hz refresh rate. The centre of the screen was positioned at eye level, 60-70 cm from the participants' eyes. For auditory stimuli an LTB Magnum 5.1 AC97 Headphone set was used.

D. Stimuli

The visual stimulus was based only on one scene, see Figure 2. Two animations, at a resolution of 800×600 , were rendered for the scene, see Figure 3. The animations were rendered using our own implementation of path tracing [Kaj86]. All scenes were static with only frontal camera movement and no rotation relative to the motion path. For each of them a curved motion path with the oscillating motion [46] of the camera along the vertical axis was used, see Figure 4. The stride length in walking and running animations was 0.8m and 1.5m respectively. The young subjects' average normal walking speed of 1.425 m/s [47] was used for the walking condition. For the running condition a speed of 4 m/s was used. All videos were compressed using XviD MPEG-4 Codec (single-pass encoding, target quantizer: 3.00). The animations were divided into three *walking animations* and two *running animations* in both ways, each lasting for five seconds.

For audio, we chose foot steps as the animation-related sound effect. Sound effect of foot steps, made by leather soled shoes against a firm tiled floor was used as the auditory stimulus in both camera conditions (walking and running). To synchronise the sound effects with the animation, the length of the silence between the ON signals was varied. The amount of echo in the effect was adjusted accordingly to the nature of the scene and did not change during the animations. Sounds were produced uncompressed, using two channels (stereo), sample rate 44100Hz and bit rate of 1411kbps. We did not play any background music in order to avoid any subjective side effects such as happiness, sadness, anxiety, excitement, boredom, etc. If any of those factors are present, the perception and therefore the results could be affected [45].

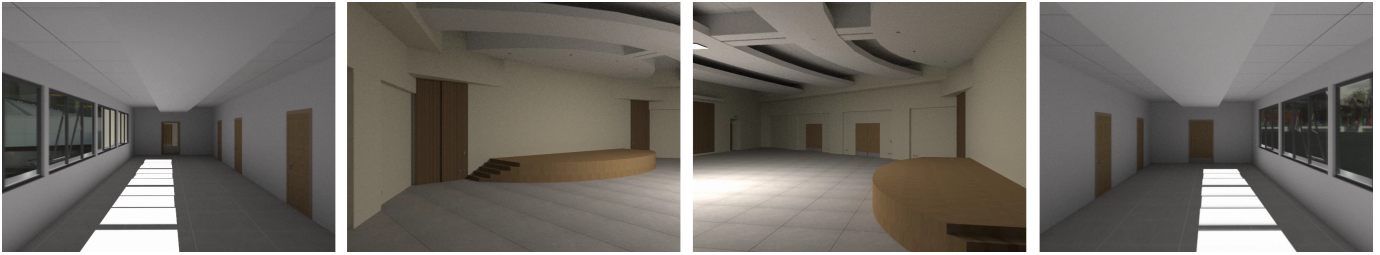


Fig. 2. Four frames taken from the walk-through animations. The first two frames are from the animations with camera moving from the corridor to the conference hall, and last two from the animations where the camera is moving from the conference hall to the corridor.

For the audio-visual presentation we developed a framework with support for frame rate and audio control. All results from each trial were saved in separate text files.

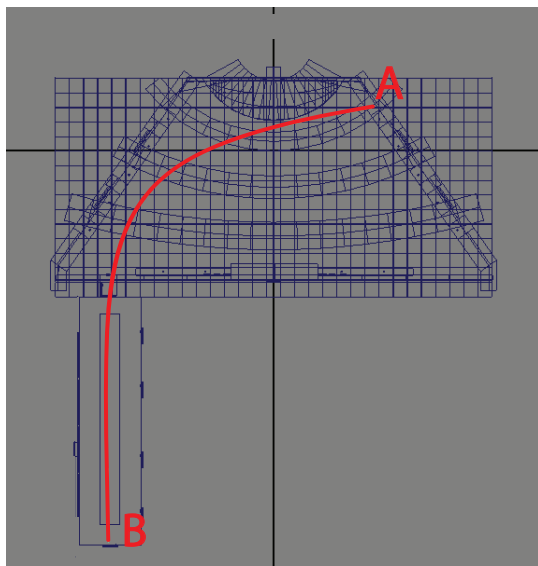


Fig. 3. Camera path used for the animations. Two way of movement were from A to B and from B to A.

E. Procedure

Prior to the experiment each participant was asked to read and sign the consent form, in which they agreed to voluntarily and anonymously participate in the experiment and being able to withdraw from it at any time. Participants were also given a questionnaire to fill in. After this they were presented with the instructions, followed by two sample animation pairs played at 10 and 60fps. They were told that these are the worst and the best cases respectively, but not what frame rates the animations were. In the instructions they were further explained what is frame rate and that they are going to watch 22 pairs of animations whose smoothness they will have to evaluate.

The randomly ordered animation pairs were presented sequentially, Figure 5. Each animation was preceded by a grey box (RGB: 0.3, 0.3, 0.3) lasting for two seconds. The length of each animation was five seconds. After each pair the A and B boxes were shown on the screen. Participants were

instructed to choose the smoother animation by clicking on one of the two boxes, after which the next cycle would start automatically. Each trial lasted for about six minutes. After the experiments they were debriefed on the nature of the study.

IV. RESULTS

In order to test our hypotheses, the data was analysed using descriptive statistics and a non-parametric Chi-square test [48], [49].

A. Walking vs running - visual smoothness perception

Our first research hypothesis was that the speed of the camera movement will affect the smoothness perception. The null hypothesis was that camera movement speed will have no effect on visual smoothness perception.

We tested this hypothesis comparing the walking (Walk) and running (Run) animations played at 60fps, which is our gold standard. We tested separately *noAudio* (silent) condition and *Audio* condition. This had a single independent variable - camera movement speed.

Tabular data with observed and expected frequencies are given in Table I.

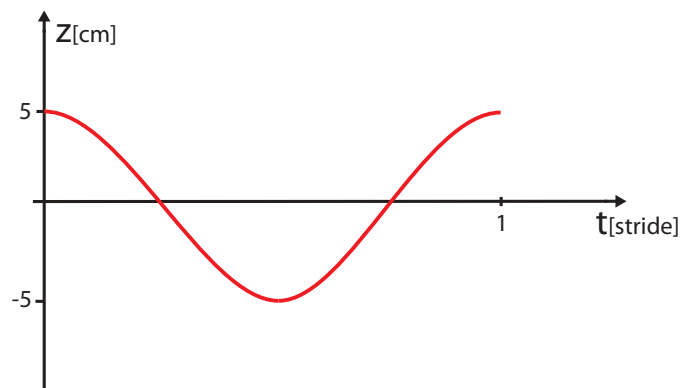


Fig. 4. Oscillating camera motion along the vertical (z) axis

The Chi-square test of the relationship between Running and Walking animations for noAudio condition produced $\chi^2(1)=5.444$, $df=1$, $p=.020$. As the analysis revealed a statistically significant association, the null hypothesis was rejected in favour of the research hypothesis that the speed of the camera



Fig. 5. The experimental procedure. From left to right: grey box, first animation, grey box, second animation and A/B evaluation screen.

movement affects the animation smoothness perception. Participants preferred the walking rather than running animation.

The same test for Audio condition produced $\chi^2(1)=1.778$, $df=1$, $p=.182$, and thus the null hypothesis cannot be rejected. This indicates that audio might affect the perception of the animation smoothness, which we investigate further in Section IV-B.

noAudio			
	Observed N	Expected N	Residual
Run	11	18.0	-7.0
Walk	25	18.0	7.0
Total	36		

Audio			
	Observed N	Expected N	Residual
Run	14	18.0	-4.0
Walk	22	18.0	4.0
Total	36		

TABLE I

OBSERVED AND EXPECTED FREQUENCIES FOR THE RUN - WALK ANIMATION SMOOTHNESS PERCEPTION COMPARISON. TOP TABLE SHOWS FREQUENCIES FOR NOAUDIO CONDITION. BOTTOM TABLE SHOWS FREQUENCIES FOR AUDIO CONDITION.

B. Sound effect's influence on perceived smoothness threshold

In this part of the analysis we investigated the perceived smoothness threshold when watching the animations accompanied by the audio effects. The difference in the preferences between the discrepant frame rate pairs and the control group (60vs60 fps) were compared. The preference of the following frame rates were compared: 10vs60, 20vs60 and 30vs60 against 60vs60. Lower frame rates were always used with the running animation. Our null hypothesis for each test pair was that, when presented along with movement-related sound effects, animations in discrepant frame rate pairs with higher frame rates will not be perceived as smoother. Since we assume no bias, this means that each test pair will have the same preference compared to the control group.

For the analysis we used a one-tailed Chi-square test. Thus, in order to test the validity of the one-tailed hypothesis we compared corresponding means. The mean values for test

pairs (10vs60, 20vs60 and 30vs60) were 1.83, 1.52 and 1.63 respectively, where lower and upper bounds were 1 and 2 respectively. The mean value for our control group was 1.61. Since the mean value of 20vs60 condition is lower than the mean value of the control group, the null hypothesis for this pair cannot be accepted. For the 10vs60 and 30vs60 pairs the difference is in line with our research hypothesis and that we can carry on with the test.

	Mean value	p-value
10vs60	1.83	.032
20vs60	1.52	N/A*
30vs60	1.63	.5
60vs60	1.61	N/A

TABLE II

MEAN AND P VALUES FOR AUDIO CONDITION. P-VALUE IS GIVEN FOR DIFFERENCE IN PREFERENCE BETWEEN THE TEST PAIRS AND 60VS60 CONDITION. LOWER AND UPPER BOUNDS WERE 1 (FIRST ANIMATION PREFERRED) AND 2 (SECOND ANIMATION PREFERRED) RESPECTIVELY. *MEAN VALUE LOWER THAN 1.61 (CONTROL GROUP)

The results show that only for 10vs60 pair there was significant difference in preference ($p=.032$), and thus the null hypothesis can be rejected. For 30vs60 ($p=.5$) pairs there was no significant difference in preference comparing to the 60vs60 control group. Hence, the null hypothesis in this case cannot be rejected. These results appear to show that the perceived smoothness threshold when watching the animations with movement-related sound effects is somewhere between the 10 and 20 fps. This further indicates that animations of a walking person, rendered at 60 frames per second, when accompanied by the movement-related sound effects, were not perceived significantly smoother than the same animation rendered at 20 and 30 fps.

V. CONCLUSION AND FUTURE WORK

In this paper we investigated the effect of camera movement speed and movement-related sound effects on the perception of animation smoothness. The results show that walking animations are perceived as smoother when presented without any sound effects. We also showed that movement-related sound effects do indeed improve the perceived smoothness quality of the animations.

Such results indicate a step forward in our understanding of auditory-visual cross-modal interaction and its utilisation in computer graphics. Once understood and harnessed correctly this could be used for interactive systems, such as high-fidelity rendering for serious game applications, when the camera movement speed increases, possibly leading to an increase in the computational requirements and a drop in the produced frame rate. In such cases, additional sound effects could be introduced in order to maintain the same perceptual smoothness quality.

In the future, we will compare the audio-visual content with lower frame rates with ones presented at higher frame rates. We would also like to investigate the same effect in interactive scenarios with and without user tasks. Such results also entail the possibility of building decision-theoretic systems [50] based on cross-modal effects that ensure a constant perceived frame rate rather than the commonly used fixed frame rate [51].

ACKNOWLEDGMENT

The authors would like to thank all the participants from University of Warwick who volunteered for the experiments. We would also like to thank Piotr Dubla for helping us with the framework development.

REFERENCES

- [1] A. Chalmers, D. Howard, and C. Moir, "Real virtuality: A step change from virtual reality," in *Spring Conference on Computer Graphics 2009*, H. Hauser, Ed. ACM SIGGRAPH Press, 2009, pp. 15–22.
- [2] A. Chalmers, K. Debattista, and B. Ramic-Brkic, "Towards high-fidelity multi-sensory virtual environments," *The Visual Computer*, vol. 25, no. 12, pp. 1101–1108, 2009.
- [3] J. W. Gebhard and G. H. Mowbray, "On discriminating the rate of visual flicker and auditory flutter," *Am J Psychol*, vol. 72, pp. 521–529, Dec 1959.
- [4] T. Shipley, "Auditory flutter-driving of visual flicker," *Science*, vol. 145, pp. 1328–1330, Sep 1964.
- [5] L. Shams, Y. Kamitani, and S. Shimojo, "What you see is what you hear," *Nature*, vol. 408, pp. 788+, 2000.
- [6] —, "Visual illusion induced by sound," *Brain Res Cogn Brain Res*, vol. 14, no. 1, pp. 147–152, Jun 2002.
- [7] G. H. Recanzone, "Auditory influences on visual temporal rate perception," *Journal of neurophysiology*, vol. 89, pp. 1078–1093, Feb 2003.
- [8] S. Morein-Zamir, S. Soto-Faraco, and A. Kingstone, "Auditory capture of vision: examining temporal ventriloquism," *Brain Res Cogn Brain Res*, vol. 17, no. 1, pp. 154–163, Jun 2003.
- [9] L. Shams, Y. Kamitani, and S. Shimojo, "Modulations of visual perception by sound. in the handbook of multisensory processes (eds. calvert, g.a., spence, c. and stein, b.e.)," pp. 27–33, 2004.
- [10] G. Mastoropoulou, "The effect of audio on the visual perception of high-fidelity animated 3d computer graphics," PhD in Computer Science, University of Bristol, 2006.
- [11] G. Mastoropoulou, K. Debattista, A. Chalmers, and T. 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*. New York, NY, USA: ACM Press, 2005, pp. 9–15.
- [12] V. Hulusic, G. Czanner, K. Debattista, E. Sikudova, P. Dubla, and A. Chalmers, "Investigation of the beat rate effect on frame rate for animated content," in *Spring Conference on Computer Graphics 2009*, H. Hauser, Ed. ACM SIGGRAPH Press, 2009, pp. 167–174.
- [13] W. James, *The principles of psychology*. New York: Holt, 1890.
- [14] A. Mack and I. Rock, *Inattentive Blindness*. The MIT Press, 1998.
- [15] D. Simons and C. Chabris, "Gorillas in our midst: sustained inattentive blindness for dynamic events," *perception*, vol. 28, pp. 1059–1074, 1999.
- [16] M. C. Kelly and A. I. Tew, "The continuity illusion in virtual auditory space," in *in proc. of AES 112th Convention*, Munich, Germany, May 2002.
- [17] B. C. Moore, *An Introduction to the Psychology of Hearing*, 2nd ed. Academic Press, 1982.
- [18] G. W. Humphreys and V. Bruce, *Visual Cognition: Computational, Experimental and Neuropsychological Perspectives*. East Sussex, BN3 2FA, UK: Lawrence Erlbaum Associates Ltd, 1989.
- [19] I. P. Howard and W. B. Templeton, *Human spatial orientation [by] I.P. Howard and W.B. Templeton*. Wiley, London, New York., 1966.
- [20] H. McGurk and J. Macdonald, "Hearing lips and seeing voices," *Nature*, vol. 264, no. 5588, pp. 746–748, December 1976.
- [21] B. Ramic-Brkic, A. Chalmers, K. Boulanger, S. Pattanaik, and J. Covington, "Cross-modal affects of smell on the real-time rendering of grass," in *Spring Conference on Computer Graphics 2009*, H. Hauser, Ed. ACM SIGGRAPH Press, 2009, pp. 175–179.
- [22] M. Keetels and J. Vroomen, "Tactile-visual temporal ventriloquism: no effect of spatial disparity," *Percept Psychophys*, vol. 70, no. 5, pp. 765–771, Jul 2008.
- [23] R. B. Welch and D. H. Warren, "Immediate perceptual response to intersensory discrepancy," *Psychological bulletin*, vol. 88, no. 3, pp. 638–667, November 1980.
- [24] C. S. Choe, R. B. Welch, R. M. Gilford, and J. F. Juola, "The 'ventriloquist effect': Visual dominance or response bias?" *Perception and Psychophysics*, vol. 18, no. 1, pp. 55–60, 1975.
- [25] J. Vroomen, P. Bertelson, and B. d. Gelder, "A visual influence in the discrimination of auditory location," 1998.
- [26] J. Vroomen and B. de Gelder, "Perceptual effects of cross-modal stimulation: Ventriloquism and the freezing phenomenon. in the handbook of multisensory processes (eds. calvert, g.a., spence, c. and stein, b.e.)," pp. 140–150, 2004.
- [27] Y. Wada, N. Kitagawa, and K. Noguchi, "Audio-visual integration in temporal perception," *Int J Psychophysiol*, vol. 50, no. 1-2, pp. 117–124, October 2003.
- [28] R. Sekuler, A. B. Sekuler, and R. Lau, "Sound alters visual motion perception," *Nature*, vol. 385, no. 6614, p. 308, January 1997. [Online]. Available: <http://dx.doi.org/10.1038/385308a0>
- [29] Y. Kamitani and S. Shimojo, "Sound-induced visual 'rabbit'," *J. Vis.*, vol. 1, no. 3, pp. 478–478, 12 2001. [Online]. Available: <http://journalofvision.org/1/3/478/>
- [30] S. Getzmann, "The effect of brief auditory stimuli on visual apparent motion," *Perception*, vol. 36, no. 7, pp. 1089–1103, 2007.
- [31] D. A. Allport, B. Antonis, and P. Reynolds, "On the division of attention: a disproof of the single channel hypothesis," *Q J Exp Psychol*, vol. 24, no. 2, pp. 225–235, May 1972.
- [32] A. M. Bonnel and E. R. Hafter, "Divided attention between simultaneous auditory and visual signals," *Percept Psychophys*, vol. 60, no. 2, pp. 179–190, Feb 1998.
- [33] D. Alais, C. Morrone, and D. Burr, "Separate attentional resources for vision and audition," *Proc Biol Sci*, vol. 273, no. 1592, pp. 1339–1345, Jun 2006.
- [34] D. Burr and D. Alais, "Combining visual and auditory information," *Prog Brain Res*, vol. 155, pp. 243–258, 2006.
- [35] J. Duncan, S. Martens, and R. Ward, "Restricted attentional capacity within but not between sensory modalities," *Nature*, vol. 387, no. 6635, pp. 808–810, Jun 1997.
- [36] D. W. Massaro and D. S. Warner, "Dividing attention between auditory and visual perception," *Perception & Psychophysics*, vol. 21, no. 6, pp. 569–574, 1977.
- [37] A. Larsen, W. McIlhagga, J. Baert, and C. Bundesen, "Seeing or hearing? perceptual independence, modality confusions, and crossmodal congruity effects with focused and divided attention," *Percept Psychophys*, vol. 65, no. 4, pp. 568–574, May 2003.
- [38] H. Yee, S. Pattanaik, and D. P. Greenberg, "Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments," *ACM Trans. Graph.*, vol. 20, no. 1, pp. 39–65, 2001. [Online]. Available: <http://pdiff.sourceforge.net/ypg01.pdf>
- [39] 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 Techniques*. Leuven, Belgium: Eurographics Association, 2003, pp. 270–280.
- [40] C. Kayser, C. I. Petkov, M. Lippert, and N. K. Logothetis, "Mechanisms for allocating auditory attention: An auditory saliency map," *Current*

Biology, vol. 15, no. 21, pp. 1943–1947, November 2005. [Online]. Available: <http://dx.doi.org/10.1016/j.cub.2005.09.040>

- [41] T. Moeck, N. Bonneel, N. Tsingos, G. Drettakis, I. Viaud-Delmon, and D. Alloza, “Progressive perceptual audio rendering of complex scenes,” in *13D '07: Proceedings of the 2007 symposium on Interactive 3D graphics and games*. New York, NY, USA: ACM, 2007, pp. 189–196.
- [42] N. Tsingos, E. Gallo, and G. Drettakis, “Perceptual audio rendering of complex virtual environments,” *ACM Trans. Graph.*, vol. 23, no. 3, pp. 249–258, 2004.
- [43] G. Mastoropoulou, K. Debattista, A. Chalmers, and T. 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*. New York, NY, USA: ACM Press, 2005, pp. 363–369.
- [44] V. Hulusic, M. Aranha, and A. Chalmers, “The influence of cross-modal interaction on perceived rendering quality thresholds,” in *WSCG 2008 Full Papers Proceedings*, V. Skala, Ed., 2008, pp. 41–48.
- [45] G. Mastoropoulou and A. Chalmers, “The effect of music on the perception of display rate and duration of animated sequences: An experimental study,” in *TPCG '04: Proceedings of the Theory and Practice of Computer Graphics 2004 (TPCG'04)*. Washington, DC, USA: IEEE Computer Society, 2004, pp. 128–134.
- [46] A. Lecuyer, J.-M. Burkhardt, J.-M. Henaff, and S. Donikian, “Camera motions improve the sensation of walking in virtual environments,” in *VR '06: Proceedings of the IEEE conference on Virtual Reality*. Washington, DC, USA: IEEE Computer Society, 2006, pp. 11–18.
- [47] M. Arif, Y. Ohtaki, R. Nagatomi, and I. H., “Estimation of the effect of cadence on gait stability in young and elderly people using approximate entropy technique,” *MEASUREMENT SCIENCE REVIEW*, vol. 4, pp. 29–40, 2004.
- [48] D. C. Howell, *Statistical Methods for Psychology*, 6th ed. Thomson Wadsworth, 2007.
- [49] *Exploring Psychological Research Methods*. The Open University, 2007.
- [50] R. Dumont, F. Pellacini, and J. A. Ferwerda, “Perceptually-driven decision theory for interactive realistic rendering,” *ACM Trans. Graph.*, vol. 22, no. 2, pp. 152–181, 2003.
- [51] T. A. Funkhouser and C. H. Séquin, “Adaptive display algorithm for interactive frame rates during visualization of complex virtual environments,” in *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM, 1993, pp. 247–254.