Self-induced Footsteps Sounds in Virtual Reality: Latency, Recognition, Quality and Presence

Rolf Nordahl
Medialogy, Aalborg University Copenhagen
2750 Ballerup, Denmark
rn@media.aau.dk

Abstract

In this paper we describe the results of experiments whose goal is to investigate the effect of enhancing a virtual reality experience with the sound of synthetic footsteps. Results show that the sense of presence is enhanced when the sound of one's own motion is added. Furthermore, the experiments show that the threshold for detection of latency between motion and sound is raised when visual stimuli is introduced.

1. Introduction

Through the recent years some presence studies have focused on whether the addition of auditory cues in Virtual Environments (VE) and Virtual Reality (VR) could lead to measurable enhancements in participators feeling of presence. Most of the results of previous research have been focusing of sound delivery methods [1,2,4], sound quantity [3,4] and quality of visual versus auditory information [3,5]. To our knowledge, the effect of self induced sounds to enhance sense of presence has not been investigated yet. In this paper, we are interested in investigating if enhancing the VR experience with the sound of the subjects' own footsteps enhances sense of presence.

We designed a real-time footstep synthesizer, controlled by the subjects by using a set of sandals embedded with pressure sensitive sandals. By navigating in the environment, the user controls the synthetic sounds.

2. Designing synthetic footsteps

The footstep sound synthesizer works in real-time under the Max/MSP platform[1]. Footsteps recorded on seven different surfaces were obtained from the Hollywood Edge Sound Effects library[2]. The surfaces used were metal, wood, grass, bricks, tiles, gravel and snow.

The sounds were analyzed using the Fast Fourier Transform (FFT), and the main resonances were extracted from the spectrum. .Such resonances were used to build a modal synthesizer [6, 7]. To trigger the synthetic footsteps, the users were asked to wear a pair of sandals embedded with pressure sensitive sensors placed one in each heel.

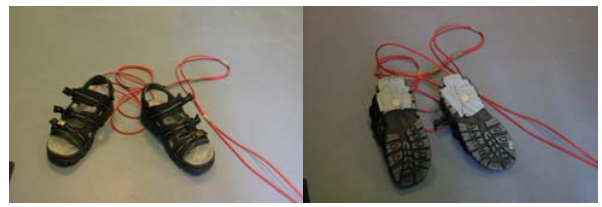

Figure 1: The shoe controller developed for the experiments.

When the subject walked around, the heel of the sandal would come into contact with the floor, thereby trigging the sensors. Through the use of a microprocessor, the corresponding pressure value was converted into an input parameter which was read by the real-time sound synthesizer Max/MSP. The sensors were connected to the microprocessor through wires, as shown in Figure 1, and the microprocessor was connected to a laptop PC.

3. Experimental setup

The main goal of the experiments was to test if the sound of one's own footsteps enhanced the sense of presence. Two different experiments were performed.

In the first experiment only the auditory feedback was provided. Subjects were asked to wear the sandals enhanced with sensors and a pair of headphones without being told the purpose of either. In this experiment latency-perception and auditory recognition of the floor surface were tested.

In the second experiment subjects were exposed to a VE provided through a HMD. The subjects were divided into 2 groups, one only exposed to visual feedback while the other was exposed to bimodal (audio-visual) feedback. The visual material was a reconstruction of the Prague technical museum developed as a part of the BENOGO-project. 16 subjects participated in both experiments.

4. Sound recognition, quality and evaluation

In order to test the shoe controller on different synthesized surfaces, to understand how their quality and appropriateness were perceived, a testing scenario was designed.

The seven different synthesized surfaces were played in random order, and subjects were asked to recognize the surface and judge the quality of the sound in a scale from 1 to 5 (unimodal case) or judge the appropriateness of the sound in the displayed scenario (bimodal case).

[1] www.cycling74.com
[2] www.hollywoodedge.com

‘	Unimodal		Bimodal	
	Recognition	Quality	Recognition	Appropriate
Metal	93,3%	3,9	70%	2,05
Wood	37,5%	3,7	60%	2,95
Grass	18,8%	3,25	25%	1,63
Bricks	37,5%	3,81	70%	4
Tiles	6,25%	3,78	60%	3,8
Gravel	93,75%	3,78	100%	1,6
Snow	37,5%	2,53	35%	1,47

Table 1. Result of the sound identification test, in the unimodal and bimodal condition

Table 1 shows the results of the sound identification test. Notice how subjects could easily recognize the metallic surface, more likely in the unimodal (93,3 %) than in the bimodal case (70%). Notice also the high recognition factor of gravel (93,75% in unimodal and 100% in bimodal). Hard surfaces such as wood, tiles and bricks were harder to identify, and often confused among each others. Consistent with the fact that the floor of the technical museum was made of bricks, this sound was considered most appropriate. Notice also how the recognition of such surface significantly increases when visual feedback is provided (37,5%) versus (70%).

5. Latency perception

The goal of this experiment was to test the level of acceptance of latency in VR. Subjects were asked to walk around and inform the facilitator when they perceived a delay between the step and the corresponding sound. While the subjects were walking, the facilitator was increasing the delay between the steps and the corresponding sounds, by a factor of 5 ms. The test was performed both without and with visual feedback.

	Uni-modal		Bi-modal			
	M	SD	M	SD	T	p
Latency perceived	41.7ms	5.8ms	60.9ms	20.7ms	-3.5	0.002

Table 2: Result of the latency test in the unimodal and bimodal condition.

The results of the latency perception test are shown in Table 2. Notice how the perceived latency is significantly higher in the audio-visual condition (M= 60.9 ms, SD=20.7 ms) rather than in the auditory only condition (M=41.7 ms, SD=5.8 ms). This result is most likely due to the fact that the attention of the subjects in the bimodal condition was mostly focused on the visual rather that on the auditory feedback, which was clearly not the case in the unimodal condition.

6. Presence test

In order to measure the subjective feeling of presence in unimodal (visual) and bimodal (audio-visual) case, the Swedish Viewer User Presence questionnaire [8] was chosen. 20 participants were randomly assigned either the visual (n 9, one female), or the audio-visual condition (n 10, one female). In both conditions subjects were asked to wear the HMD, headphones and sandals. In order to facilitate the self-motion, subjects were asked to count the number of airplanes and cars they could identify in the virtual space.

	Unim		Bim		T	P
	M	SD	M	SD		
PRESENCE	4.64	0.63	5.35	0.39	-2.88	0.012
ENJOYMENT	5.44	1.01	6.01	1.22	-1.27	0.2
EXTERNAL AWARENESS	-1.1667	2.97	0.2	2.33	-1.1	0.28
SIMULATOR SICKNESS	1.35	0.77	1.31	0.29	0.1	0.9

Table 3: Results of the SVUP Presence Questionnaire

Results displayed in Table 3 show that the sense of presence is significantly higher in the bimodal (M=5.35, STD=0.39) than in the unimodal case (M=4.64, SD=0.63)(P=0.012, t=-2.88).

7. Conclusions

Results obtained show that the sense of presence is significantly enhanced when self sound is added to the VR environment. However it should be noted that no condition with other kinds of sound was tested since the original real scenario did not contain any sounds, other than a very distant noise from a fan. In future tests such conditions will be added. Furthermore, tests involving both conditions with HMD and CAVE setups will be used.

References

[1] R.L. Storms and M.J. Zyda. Interactions in perceived quality of auditory-visual displays. Presence -Teleoperators an Virtual Environments, vol. 9, n. 6, pp. 557-580, 2000.

[2] Sanders, Richard D. and Scorgie, Jr. Mark A., “The Effect of Sound Delivery Methods on A User’s Sense of Presence in A Virtual Environment”, Thesis, Naval Postgraduate School, Monterey, California, 2002

[3] P. Chueng and P. Marsden. Designing Auditory Spaces: The Role of Expectation, HCI International, 2003.

[4] S. Serafin and G. Serafin. Sound design to enhance presence in photorealistic virtual reality. Proc. ICAD 2004.

[5] J. Freeman and J. Lessiter. Here, there and everywhere: The. effect of multichannel audio on presence. Proc. ICAD,. 2001

[6] L. Savioja et al. Creating Interactive Virtual Acoustic Environments. Journal of the Acoustical Engineering Society, vol. 47, no.9, pp.675-705, 1999.

[7] Cook. Real sound synthesis for interactive applications. AK Peters, 2002

[8] P. Larsson et al. The actor-observer effect in virtual reality presentations. Cyberpsychol Behav. 2001 Apr;4(2):239-46.