Aalborg Universitet

Conflicting audio-haptic feedback in physically based simulation of walking sounds

Turchet, Luca; Serafin, Stefania; Dimitrov, Smilen; Nordahl, Rolf

Published in:
Lecture Notes in Computer Science

DOI (link to publication from Publisher):
10.1007/978-3-642-15841-4_11

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

Conflicting Audio-haptic Feedback in Physically Based Simulation of Walking Sounds

Luca Turchet, Stefania Serafin, Smilen Dimitrov, and Rolf Nordahl

Medialogy, Aalborg University Copenhagen
Lautrupvang 15, 2750 Ballerup, DK
{tur,sts,sd,rn}@media.aau.dk

Abstract. We describe an audio-haptic experiment conducted using a system which simulates in real-time the auditory and haptic sensation of walking on different surfaces. The system is based on physical models, that drive both the haptic and audio synthesizers, and a pair of shoes enhanced with sensors and actuators. Such experiment was run to examine the ability of subjects to recognize the different surfaces with both coherent and incoherent audio-haptic stimuli. Results show that in this kind of tasks the auditory modality is dominant on the haptic one.

1 Introduction

While several studies have investigated the interaction between touch and audition in hand based interactions, to our knowledge, the interaction of auditory and haptic feedback in foot based devices is still an unexplored topic.

A notable exception is the work of Giordano et al., who showed that the feet were also effective at probing the world with discriminative touch, with and without access to auditory information. Their results suggested that integration of foot-haptic and auditory information does follow simple integration rules [1].

In previous research, we described a system able to simulate the auditory and haptic sensation of walking on different materials and presented the results of a preliminary surface recognition experiment [2]. This experiment was conducted under three different conditions: auditory feedback, haptic feedback, and audio-haptic feedback. By presenting the stimuli to the participants passively sitting in a chair, we introduced a high degree of control on the stimulation. However, this method of delivery is highly contrived since it eliminates the tight sensorimotor coupling that is natural during walking and foot interaction. It is true for the auditory channel, but even more so for the haptic channel. In spite of these drastically constrained conditions, performance was surprisingly good.

In particular, the results indicated that subjects were able to recognize most of the stimuli in the audition only condition, and some of the material properties such as hardness in the haptics only condition. Nevertheless, the combination of auditory and haptic cues did not significantly improve recognition.

In a successive research we extended that work improving the developed technology which allowed subjects to walk in a controlled laboratory, where their

R. Nordahl et al. (Eds.): HAID 2010, LNCS 6306, pp. 97–106, 2010.

steps were tracked and used to drive the simulation [4]. Overall, results showed that subjects were able to recognize most of the synthesized surfaces with high accuracy. Results moreover confirmed that auditory modality is dominant on the haptic modality and that the haptic task was more difficult than the other two. Indeed such results showed that subjects performed the recognition task better when using auditory feedback versus haptic feedback, and that the combination of auditory and haptic feedback only in some conditions significantly enhanced the recognition.

Starting from those results, in this paper we investigate in a deeper way the role of dominance of the two modalities involved by means of a preliminary discrimination experiment. In particular, while in previous research we focused on providing coherent stimuli in the auditory and haptic modality, here we provide conflicting stimuli, to understand which modality is dominant.

The results presented in this paper are part of the Natural Interactive Walking (NIW) FET-Open project[1], whose goal is to provide closed-loop interaction paradigms enabling the transfer of skills that have been previously learned in everyday tasks associated to walking. In the NIW project, several walking scenarios are simulated in a multimodal context, where especially audition and haptic feedback play an important role.

2 Simulation Hardware and Software

We developed a system which simulates in real-time the auditory and haptic sensation of walking on different surfaces. A schematic representation of this system is shown in Figure 1. In order to provide both audio and haptic feedback, haptic shoes enhanced with pressure sensors have been developed. The way pressure sensors and actuators are embedded in the sandals can be seen in Figure 2, and a picture of a user wearing the shoes is shown in Figure 3. A complete description of such system and of all its components is given elsewhere in detail [5].

The hardware allows to control in real-time of a sound synthesis engine based on physical models. Such engine is illustrated in our previous research [3,7,6]. The same physical models have been used to drive the haptic and the audio synthesis.

As mentioned in section 1, the system has been evaluated by using it both offline and interactively. The complete results of this evaluation are described in [2,4].

3 Experiment

We conducted an experiment whose goal was to investigate the role of dominance of the audio and haptic modalities during the use of our walking system. Subjects were asked to interact with the system and to recognize the different walking sounds and vibrations they were exposed to.

[1] http://www.niwproject.eu/

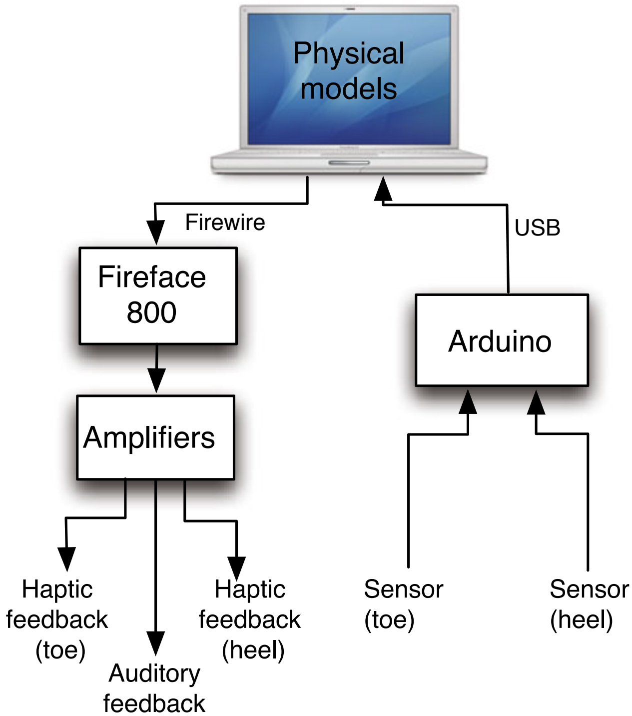

Fig. 1. Diagram illustrating the different hardware components of the system, together with their connections to the PC. The representation is for one shoe.

The experiment consisted of both coherent and incoherent audio-haptic stimuli. In presence of coherent stimuli the same surface material was presented both at audio and haptic level. Instead the provided incoherent stimuli consisted of different surface materials; in particular when at audio level a solid surface was presented, at haptic level an aggregate surface was modeled, and viceversa.

One of our hypotheses was that the audio modality would have dominated the haptic one. Another was that the recognition would have slightly improved using coherent stimuli rather than the incoherent ones. Similarly we hypothesized higher evaluations in terms of realism and quality in presence of coherent stimuli.

3.1 Participants

Ten participants, 7 male and 3 female, aged between 20 and 38 (mean = 25.81, standard deviation = 5.77), were involved in the experiment.

All participants reported normal hearing conditions and all of them were naive with respect to the experimental setup and to the purpose of the experiment.

The participants took on average about 11 minutes to complete the experiment.

Fig. 2. A picture of one pressure sensor and two actuators embedded in the shoes

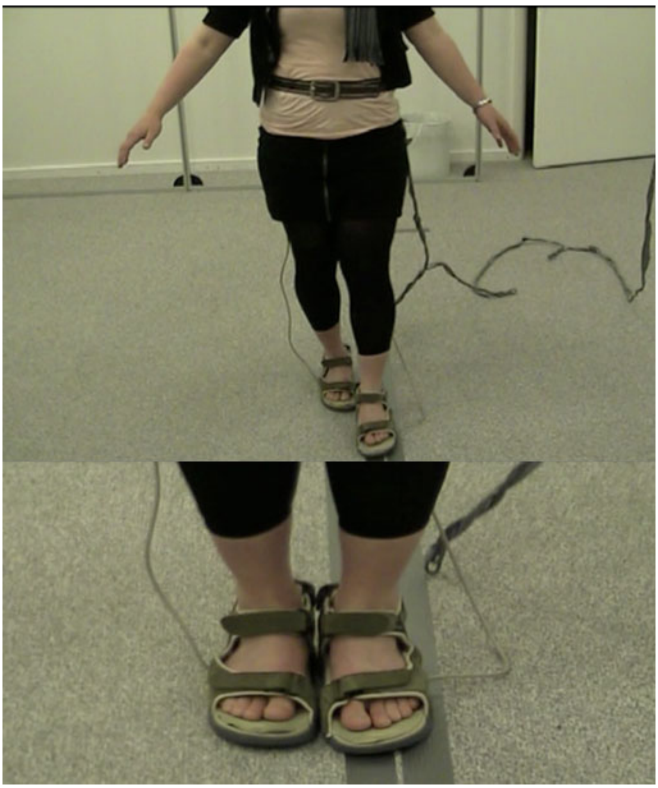

Fig. 3. Top: A user wearing the sandals enhanced with sensors and actuators. Bottom: the sandals.

3.2 Setup

The experiment was carried out in an acoustically isolated laboratory. The walking area was approximately 18 square meters, delimited by the walls of the laboratory.

The setup consisted of the pair of sandals mentioned in section 2, an Arduino board, a Fireface soundcard, a laptop and a set of headphones[2]. In order to facilitate the navigation of the subjects, the wires coming out from the shoes in all setups, as well as the wires connecting the headphones to the soundcard, were linked to a bumbag or to snaplinks attached to trousers.

3.3 Task

During the experiment participants were asked to wear the pair of sandals and the headphones described in sections 2 and 3.2, and to walk in the laboratory.

During the act of walking they listened simultaneously to footsteps sounds and vibrations on a different surface according to the stimulus presented. The task consisted of answering, by writing on a paper, the following three questions after the presentation of the stimulus:

1. Which surface do you think you are walking on? For each stimulus choose a material.
2. How close to real life is the sound in comparison with the surface you think it is? Evaluate the degree of realism on a scale from 1 to 7 (1=low realism, 7=high realism).
3. Evaluate the quality of the sound on a scale from 1 to 7 (1=low quality, 7=high quality).

As opposed to our previous research, participants were not provided with a forced list of possible choices. This was due to the fact that we wanted subjects to be somehow creative in their recognition of the surface, without guessing from a predefined list.

Subjects were informed that they could use the interactive system as much as they wanted before giving an answer. They were also told that they could choose the same material more than one time. When passed to the next stimulus they could not change the answer to the previous stimuli.

At the conclusion of the experiment, participants were asked some questions concerning the naturalness of the interaction with the system and to comment on its usability and possible integration in a virtual reality environment. In particular the questionnaire was the following:

– Imagine that this is part of a system used to navigate in a computer game, answer to the following questions:
    1. How natural is the interaction? Evaluate on a scale from 1 to 7 (1=little natural, 7=very natural)

[2] Sennheiser HD 600, http://www.sennheiser.com

2. How normal do you feel during the act of walking? Evaluate on a scale from 1 to 7 (1=little normal, 7=very normal)
3. How constrained do you feel during the act of walking? Evaluate on a scale from 1 to 7 (1=little constrained, 7=very constrained)

In addition they were also given the opportunity to leave an open comment on their experience interacting with the system.

3.4 Experimental Plan

Participants were exposed to 12 trials consisting of 4 coherent stimuli and 8 incoherent stimuli. The 12 audio-haptic stimuli were presented once in randomized order.

The modeled surfaces were 4 (2 solid and 2 aggregate): wood, metal, snow and gravel. In presence of incoherent stimuli the conflict was rendered providing on the one hand one of the solid surfaces by means of auditory feedback, while at haptic level one of the aggregate surfaces was presented. On the other hand, analogously, another set of incoherent stimuli consisted of aggregate surfaces at auditory level while a solid surface was presented by means of haptic feedback.

4 Results and Discussion

Table 1 shows the confusion matrix resulting from the experiment. In such table are illustrated the participants answers gathered according to three material categories: solid, aggregate and liquid.

The first noticeable element emerging from the table is that none of the participants classified as liquid the simulated surfaces. Moreover it is very evident that in presence of conflicts the auditory modality is strongly dominant on the haptic one. This result is more clearly illustrated in table 2, and it seems to be confirmed by the answers concerning the names of the chosen materials (see table 3). Indeed, participants had the tendency to answer with the same names chosen when a same material was presented at auditory level both in presence of coherent and incoherent stimuli. In other words they were driven in their choice by the auditory feedback.

This does not mean that they completely ignored the haptic part of the bimodal stimulus; indeed in the comments two participants reported that they noticed that only some of the haptic stimuli where appropriate for the sound they were listening. Moreover results show that in two cases participants seemed to be driven in their choices, at least partially, by the haptic feedback; this is the case of the audio-haptic stimulus metal-snow for which one participant surprisingly gave the answer “mud” which can be considered appropriate for the haptic stimulus. Concerning the stimulus snow-wood one participant chose the answer “carpet” which seems to consider both the two components of the bimodal stimulus since the sound of snow could be interpreted as the sound on carpet in presence of an haptic feedback expressing a solid surface like wood. Both the participants, like all the others, were asked, at the conclusion of the experiment,

Table 1. Confusion matrix of experiment

Stimulus		**Answer**			
Audio	**Haptic**	**Solid**	**Aggregate**	**Liquid**	**I don't know**
Wood	Wood	7	1		2
Wood	Snow	7			3
Wood	Gravel	6	1		3
Metal	Metal	9			1
Metal	Snow	6	1		3
Metal	Gravel	8			2
Gravel	Gravel		9		1
Gravel	Wood		10		
Gravel	Metal		7		3
Snow	Snow		10		
Snow	Wood		9	1	
Snow	Metal		9		1

Table 2. Percentages of dominance of the auditory and haptic modalities in presence of incoherent stimuli

Stimulus		**Dominance**	
Audio	**Haptic**	**% Audio**	**% Haptic**
Wood	Snow	70	0
Wood	Gravel	60	0
Metal	Snow	60	10
Metal	Gravel	80	0
Gravel	Wood	100	0
Gravel	Metal	70	0
Snow	Wood	90	10
Snow	Metal	90	0

to explain their answers and they confirmed their choices. In particular all the participants were asked at the end of the experiment to classify their answers as belonging to the categories of solid, aggregate or liquid materials.

For the stimulus wood-gravel one participant reported the answer "not solid plastic" which could be addressed as haptic dominance, but the same participant chose the same answer also for the coherent stimulus wood-wood and for this reason we did not consider such answer as haptic dominance in table 2.

Regarding the percentages of "I don't know" answers, although the number is low, it is an indication of the difficulty of the proposed task. This fact was also confirmed in numerous comments left by participants who on average reported that the task was very difficult. Although the comparison between coherent and incoherent stimuli with the same auditory stimulus do not reveal any statistically significant difference (confirming from another point of view the dominance of the auditory feedback), the percentages of "I don't know" answers is on average higher for the incoherent stimuli.

Table 3. Names of the materials chosen for each audio-haptic stimulus. In bold the choices which seem to be driven by the haptic feedback.

Stimulus		**Answer**
Audio	**Haptic**	**Names of chosen materials**
Wood	Wood	wood, concrete, plastic, not solid plastic
Wood	Snow	wood, concrete, plastic, gum
Wood	Gravel	wood, concrete, plastic, not solid plastic
Metal	Metal	metal, iron, steel, wood, glass
Metal	Snow	metal, iron, steel, wood, glass, **mud**
Metal	Gravel	metal, iron, steel, wood, glass, plastic
Gravel	Gravel	gravel, little stones, sand
Gravel	Wood	gravel, little stones, sand
Gravel	Metal	gravel, little stones, sand
Snow	Snow	snow, ice, gravel, sand, leaves, not solid plastic, paper
Snow	Wood	snow, ice, gravel, sand, leaves, not solid plastic, **carpet**
Snow	Metal	snow, ice, gravel, sand, not solid plastic

Table 4. Average realism and quality scores from a seven-point Likert scale and relative standard deviation

Stimulus		**Realism**		**Quality**	
Audio	**Haptic**	μ	σ	μ	σ
Wood	Wood	2.75	1.488	3.25	1.2817
Wood	Snow	2.7143	1.496	3.2857	1.2536
Wood	Gravel	2.5714	1.8127	2.7143	1.3801
Metal	Metal	2.8889	1.6159	4	2
Metal	Snow	3.8571	1.496	3.5714	1.1339
Metal	Gravel	2.625	1.3025	2.875	1.5526
Gravel	Gravel	3.2222	1.0929	4	1
Gravel	Wood	4	0.9428	4.4	0.9661
Gravel	Metal	3.7143	1.8898	4.1429	1.4639
Snow	Snow	3.7	1.567	3.8	1.3984
Snow	Wood	3.6	1.4298	3.7	1.567
Snow	Metal	4	1.5	4.1111	1.453

Table 5. Questionnaire results. Average scores from a seven-point Likert scale and relative standard deviation.

	μ	σ
Naturalness	3.5	1.6499
Normality	4	1.5635
Constriction	4.2	1.3166

Table 4 shows the degree to which participants judged the realism and the quality of the experience. Such parameters were calculated by looking only at the answers different from "I dont know". Contrary to our hypotheses we did not find higher evaluations of these parameters for the coherent stimuli compared to the incoherent one. Surprisingly for some stimuli the evaluations are even higher for the incoherent stimuli. Anyways an in depth statistical analysis performed with the t-test revealed that all these differences are not significative.

Finally, as concerns the questionnaire conducted at the conclusion of the experiment, results in table 5 show that that subjects judged the interaction with the system not too much natural (mean = 3.5), and that they felt quite normal (mean = 4) but at the same time quite constrained (mean = 4.2) during the act of walking.

Indeed, more than one subject commented on the need of a wireless system able to convey vibrations to the shoes and sounds to the headphones set.

5 Conclusion and Future Work

In this paper, we describe an experiment conducted with a real-time footsteps synthesizer able to provide audio and haptic feedback, and which is controlled by the user during the act of walking by means of shoes embedded with sensors and actuators.

In the experiment, both coherent and incoherent audio-haptic stimuli were provided. Results confirm that auditory modality is dominant on the haptic one. This can be due to the low sensitivity of the foot when exposed to haptic signals.

The developed system is ready to be integrated in computer games and interactive installations where a user can navigate.

In future work, we indeed plan to utilize the system in multimodal environments, and include visual feedback, to understand the role of the different sensorial modalities to enhance sense of immersion and presence in scenarios where walking plays an important role.

Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme under FET-Open grant agreement 222107 NIW - Natural Interactive Walking.[3] The authors wish also to thank Amir Berrezag and Vincent Hayward who provided the shoes used in the experiments.

References

1. Giordano, B.L., Mcadams, S., Visell, Y., Cooperstock, J., Yao, H.Y., Hayward, V.: Non-visual identification of walking grounds. Journal of the Acoustical Society of America 123(5), 3412–3412 (2008)

[3] www.niwproject.eu

2. Nordahl, R., Berrezag, A., Dimitrov, S., Turchet, L., Hayward, V., Serafin, S.: Preliminary experiment combining virtual reality haptic shoes and audio synthesis. In: Proc. Eurohaptics (2010)
3. Nordahl, R., Serafin, S., Turchet, L.: Sound synthesis and evaluation of interactive footsteps for virtual reality applications. In: Proc. IEEE VR 2010 (2010)
4. Serafin, S., Turchet, L., Nordahl, R., Dimitrov, S., Berrezag, A., Hayward, V.: Identification of virtual grounds using virtual reality haptic shoes and sound synthesis. In: Proc. of Eurohaptics Symposium on Haptics and Audio-Visual Environments (2010)
5. Turchet, L., Nordahl, R., Berrezag, A., Dimitrov, S., Hayward, V., Serafin, S.: Audio-haptic physically based simulation of walking sounds. In: Proc. of IEEE International Workshop on Multimedia Signal Processing (2010)
6. Turchet, L., Nordahl, R., Serafin, S.: Examining the role of context in the recognition of walking sounds. In: Proc. of Sound and Music Computing Conference (2010)
7. Turchet, L., Serafin, S., Dimitrov, S., Nordahl, R.: Physically based sound synthesis and control of footsteps sounds. In: Proceedings of Digital Audio Effects Conference (2010)