Sound Synthesis and Evaluation of Interactive Footsteps for Virtual Reality Applications

Nordahl, Rolf; Serafin, Stefania; Turchet, Luca

Published in:
I E E E Virtual Reality Annual International Symposium

DOI (link to publication from Publisher):
10.1109/VR.2010.5444796

Publication date:
2010

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

Sound Synthesis and Evaluation of Interactive Footsteps for Virtual Reality Applications

Rolf Nordahl *   Stefania Serafin†   Luca Turchet‡

Medialogy, Aalborg University Copenhagen
Lautrupvang 15, 2750 Ballerup, DK

ABSTRACT

A system to synthesize in real-time the sound of footsteps on different materials is presented. The system is based on microphones which allow the user to interact with his own footwear. This solution distinguishes our system from previous efforts that require specific shoes enhanced with sensors. The microphones detect real footsteps sounds from users, from which the ground reaction force (GRF) is estimated. Such GRF is used to control a sound synthesis engine based on physical models. Evaluations of the system in terms of sound validity and fidelity of interaction are described.

Keywords: sound synthesis, physical models, footsteps sounds, auditory perception

Index Terms: H.5.5 [Information Systems]: Information Interfaces and Presentation—Sound and Music Computing; H.5.2 [Information Systems]: Information Interfaces and Presentation—User Interfaces

1 INTRODUCTION

The development of efficient yet accurate simulation algorithms, together with improvements in hardware technology, has boosted the research on auditory display and physically based sound models for virtual environments (VEs)[27, 23, 7].

The addition of auditory cues and their importance in enhancing the sense of immersion and presence is a recognized fact in virtual environment research and development. Most prior work in this area has focused on sound delivery methods [25, 24], sound quantity and quality of auditory versus visual information [4] and 3D sound [11, 28]. Recent studies have investigated the role of auditory cues in enhancing self-motion and presence in virtual environments [17, 15, 26].

Self-generated sounds have been often used as enhancements to VEs and first-person 3D computer games – particularly in the form of footstep sounds accompanying self-motion or the presence of other virtual humans.

Such sounds are used to produce embodiment and a sense of weight with the overall goal of heightening the sense of "realness" to the character or person. Usually such sounds are taken from sound libraries or recorded by *Foley* artists who put shoes in their hands and interact with different materials to simulate the act of walking.

Recently, several physics based algorithms have been proposed to simulate the sounds of walking. One of the pioneers in this field is Perry Cook, who proposed a collection of physically informed stochastic models (PhiSM) simulating several everyday sonic events [5]. Among such algorithms the sounds of people walking on different surfaces were simulated [6]. A similar algorithm was also proposed in [10], where physically informed models reproduced several aggregate surfaces. Procedural sound synthesis of walking has also been recently described in [9].

Previous work on interactive footwear, such as the research performed by Paradiso and coworkers [21, 3], consisted of designing shoes augmented with sensors used to control footsteps sounds. A smaller number of examples, such as recent work of Nordahl [20] and Law et al. [18], have even aimed to provide multimodal cues linked to footsteps events in such environments.

In this paper, we are particularly interested in developing a solution which requires a minimum amount of sensing technology and is shoe independent, which means that subjects can keep their own footwear while using the system. This creates several advantages from the interaction side: users do not need to wear ad-hoc designed shoes, whose wearability is decreased by the addition of several sensors.

We propose an interactive system which enables a designer to synthesize in real-time footsteps sounds of different materials. We describe the results of experiments whose goal is to test the degree of realism of the system, the ability of subjects to recognize the virtual material they are walking on, and the fidelity of interaction. The ultimate goal is to integrate this system in the simulation of multimodal virtual environments where the act of walking plays an important role.

2 THE SOUND SYNTHESIS ENGINE

We developed a physically based sound synthesis engine able to simulate the sounds of walking on different surfaces. Acoustic and vibrational signatures of locomotion are the result of more elementary physical interactions, including impacts, friction, or fracture events, between objects with certain material properties (hardness, density, etc.) and shapes. The decomposition of complex everyday sound phenomena in terms of more elementary ones has been an organizing idea in auditory display research during recent decades [12]. In our simulations, we draw a primary distinction between solid and aggregate ground surfaces, the latter being assumed to possess a granular structure, such as that of gravel.

2.1 Solid surfaces

Sonic interactions between solid surfaces have been extensively investigated, and results are available which describe the relationship between physical and perceptual parameters of objects in contact [16, 27]. Such sounds are typically short in duration, with a sharp temporal onset and relatively rapid decay.

A common approach to synthesizing such sounds is based on a lumped source-filter model, in which an impulsive excitation $s(t)$, modeling the physics of contact, is passed through a linear filter $h(t)$, modeling the response of the vibrating object as $y(t) = s(t) \star h(t)$.

Modal synthesis [1] is one widely adopted implementation of this idea. In this synthesis technique, the response model $h(t)$ is decomposed in terms of the resonant frequencies f_i of the vibrating object, also known as the modes of the object. The re-

*e-mail: rn@media.aau.dk
†e-mail:sts@media.aau.dk
‡e-mail:tur@media.aau.dk

IEEE Virtual Reality 2010
20 - 24 March, Waltham, Massachusetts, USA
978-1-4244-6236-0/10/$26.00 ©2010 IEEE

sponse is modeled as a bank of filters with impulse response $h(t) = \sum_i a_i e^{-b_i t} \sin(2\pi f_i t)$, where a_i represent the amplitudes of the modes, b_i the decay rates of the modes, and f_i the frequencies of the modes.

A footstep sound can be considered as the result of multiple micro-impact sounds between a shoe and a floor. The set of such micro-events can be thought as the result of the interaction between an *exciter* and a *resonator*. The exciter is represented by the interaction between shoe and ground. Such interaction can be either continuous, as in the case of a foot sliding across the floor, or discrete, as in the case of walking on a solid surface.

To simulate such scenarios, both an impact and friction model were implemented.

In the impact model, the excitation corresponding to each impact $s(t)$ is assumed to possess a short temporal extent and an unbiased frequency response. Such excitation consists of a discrete-time model of the force f between the two bodies, dependent on additional parameters governing the elasticity of the materials, their velocity of impact $\dot{x}$, and masses:

$$f(x,\dot{x}) = \begin{cases} -kx^{\alpha} - \lambda x^{\alpha} \dot{x} & \text{if } x > 0 \\ 0 & x \leq 0 \end{cases}$$

where α depends on the local geometry around the contact surface, and x stands for the compression of the exciter (when $x > 0$ the two objects are in contact) [2].

In the friction model we adopted a dynamic model, where the relationship between relative velocity v of the bodies in contact and friction force f is represented through a differential equation rather than static mapping. Assuming that friction results from a large number of microscopic elastic bonds, called bristles in [8], the v-to-f relationship is expressed as:

$$f(z,\dot{z},v,w) = \sigma_0 z + \sigma_1 \dot{z} + \sigma_2 v + \sigma_3 w$$

where z is the average bristle deflection, the coefficient σ_0 is the bristle stiffness, σ_1 the bristle damping, and the term $\sigma_2 v$ accounts for linear viscous friction. The fourth component $\sigma_3 w$ relates to surface roughness, and is simulated as fractal noise.

2.2 Aggregate surfaces

To synthesize aggregate surfaces, we implemented the physically informed sonic models (PhiSM) algorithm [5].

This model simulates particle interactions by using a stochastic parameterization. This means that the different particles do not have to be modeled explicitly, but only the probability that particles will create some noise is simulated. For many particle systems, this phenomenon is well taken into account by using a simple Poisson distribution, where the sound probability is constant at each time step, giving rise to an exponential probability waiting time between events.

2.3 Implementation

Using the algorithms described in the previous sections, we implemented a comprehensive collection of footstep sounds. As solid surfaces, we implemented metal and wood. In these materials, the impact model was used to simulate the act of walking, while the friction model was used to simulate the sound of creaking wood.

As aggregate surfaces, we implemented gravel, sand, snow, forest underbrush, dry leaves, pebbles and high grass. The simulated metal, wood and creaking wood surfaces were furthermore enhanced by using some reverberation. The role of reverberation is discussed in the testing section.

The sound synthesis algorithms were implemented in C++ as external libraries for the Max/MSP sound synthesis and multimedia real-time platform.[1] To enable compatibility with the Pure Data platform, [2] the algorithms were implemented using Flext. [3]

In our simulations, designers have access to a sonic palette making it possible to manipulate all such parameters, including material properties. One of the challenges in implementing the sounds of different surfaces was to find suitable combinations of parameters which provided a realistic simulation. For each simulated surface, recorded sounds were analyzed according to their combinations of events, and each subevent was simulated independently. As an example, the sound produced while walking on dry leaves is a combination of granular sounds with long duration both at low and high frequencies, and noticeable random sounds with not very high density that give to the whole sound a crunchy aspect. These different components were simulated with several stochastic models having the same density, duration, frequency and number of colliding objects.

The amplitude of the different components were also appropriately weighed, according to the same contribution present in the corresponding real sounds. Finally, a scaling factor for the subcomponents volumes gives to the whole sound an appropriate volume, in order to recreate a similar sound level which it would happen during a real footstep on each particular material.

A pilot test was run to ascertain that such a global volume plays an important role in the judgments concerning the sounds' realism and in the recognition of the surface material. Indeed, wrong settings for such a parameter can cause wrong recognitions.

3 Controlling the sound synthesis engine

The developed sound synthesis engine is controlled as following: users are asked to walk inside an area delimited by four microphones placed on the floor in a square configuration. Specifically, we used four Shure BETA 91,[4] high performance condenser microphones with a tailored frequency response designed specifically for kick drums and other bass instruments. The microphones' features made them a good candidate for the purpose of capturing footsteps sounds. In the interaction between a foot and a sole, the exciter is usually called ground reaction force (GRF), i.e., the reaction force supplied by the ground at every step. The aim of the phase of analysis has been that of extracting the GRF from the acoustic waveform. The real footsteps sounds produced are detected by the microphone, and their GRF extracted and used to control the temporal evolution of the synthetic footsteps. An example of a footstep sound and its corresponding GRF is shown in Figure 1.

4 Experiment

We conducted different experiments whose goal is to investigate the ability of subjects to recognize the different walking sounds they were exposed to. The study of human perception of locomotion sounds has addressed several properties of walking sound sources: the gender [19, 13] and posture of a walker [22], the emotions of a walker [13], the hardness and size of the shoe sole [13], and the ground material [14].

Such studies have been concerned only with recognition of sounds in an off-line scenario, where subjects were asked to listen to some sounds and classify them. In this experiment, we are interested in having subjects classify sounds both off-line, but also in an active settings, i.e., by using the developed interactive system. One of our hypotheses is that the recognition when using the interactive system is higher than in the off-line setup.

Moreover, we conducted an experiment using recorded real sounds in order to compare their recognition rate with that of the developed synthesized sounds.

[1] www.cycling74.com
[2] www.puredata.org
[3] http://puredata.info/Members/thomas/flext
[4] http://www.shure.com/

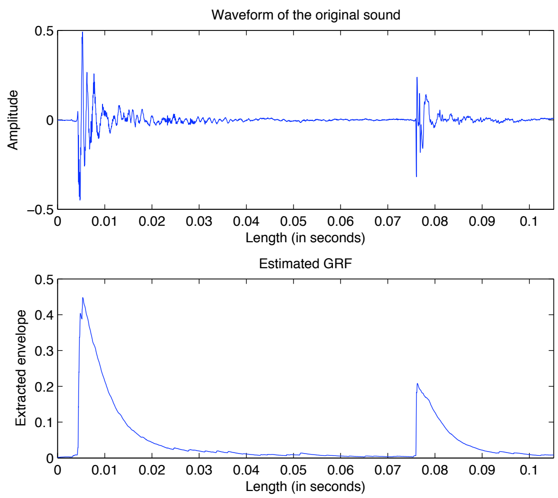

Figure 1: A footstep sound (top) and the corresponding calculated GRF (bottom).

4.1 Methods

Three kinds of experiments were conducted:

1. experiment 1: recognition of footsteps sounds generated in real time by the subjects.

2. Experiment 2: recognition of synthesized recorded footsteps sounds.

3. Experiment 3: recognition of real recorded footsteps sounds.

The sounds provided during experiment 1 were synthesized sounds generated in real time while subjects were walking using the interactive system described in the previous section.
The sounds provided during experiment 2 consisted of recordings of footsteps sounds generated by the use of the interactive system.
The sounds provided during experiment 3 consisted of recordings of real footsteps sounds on different surfaces. Such sounds were chosen among those available on the Hollywood Edge sound effects library.[5] Each sound in experiment 2 and 3, composed of several footsteps, had duration of about 7 seconds.

Participants were exposed to 26 trials in experiment 1 and 2, and 30 trials in experiment 3. During experiments 1 and 2, 13 stimuli were presented twice in randomized order. The stimuli consisted of footsteps sounds on the following surfaces: beach sand, gravel, dirt plus pebbles (like in a country road), snow (in particular deep snow), high grass, forest underbrush (a forest floor composed by dirt, leaves and branches breaking), dry leaves, wood, creaking wood and metal. To simulate room characteristics, footsteps sounds on wood, creaking wood and metal were enhanced adding a certain amount of reverberation.

In experiment 3, fifteen stimuli were presented twice in randomized order. They consisted of the previous mentioned sounds without the reverberated ones, more footsteps sounds on carpet, concrete, frozen snow, puddles and water.

4.1.1 Participants

Fourtyfive participants were divided in three groups (n = 15) to perform the three between-subjects experiments. The three groups were composed respectively of 6 men and 9 women, aged between 19 and 29 (mean=22.13,standard deviation=2.47), 8 men and 7 women, aged between 20 and 35 (mean=22.73,standard deviation=4.01), and 10 men and 5 women, aged between 20 and 29 (mean=23.13,standard deviation=2.39). All participants reported normal hearing conditions. All participants were naive with respect to the experimental setup and to the purpose of the experiment.

During experiment 1 the shoes used by subjects were sneakers, trainers, boots and other kinds of shoes with rubber soil.
The participants took in average about 24, 15 and 12 minutes for experiments 1, 2 and 3 respectively.

4.1.2 Setup

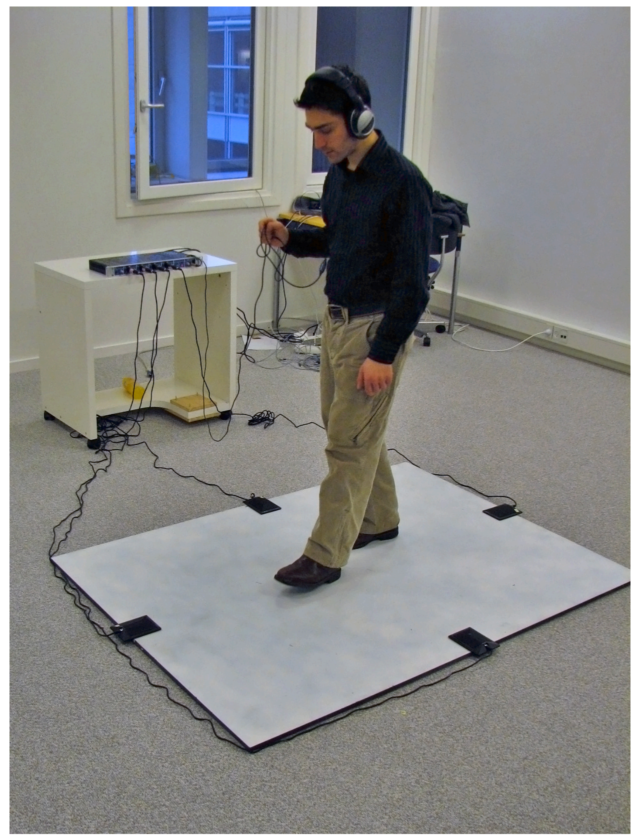

Figure 2: A subject performing experiment 1.

All experiments were carried out in an acoustically isolated laboratory where three setups were installed. The setup for experiment 1 consisted of the interactive system, and the participants were asked to use it in order to generate the footsteps sounds in real time (see Figure 2).
The setup for experiments 2 and 3 consisted of a simple graphical user interface with which the participants were asked to interact, and a spreadsheet to collect their answers.

The interface was created using the Max/MSP program and was composed only by buttons to be pressed. Each button was numbered, and by pressing it a sound was triggered and conveyed to the user by means of headphones. Users were asked to press each button according to their numerical order, and to write the corresponding answers on the spreadsheet.

[5] www.hollywoodedge.com/

	Correct answers	Wrong answers	I don't know	Realism	Quality
Beach Sand	50.	36.67	13.33	5.3	5.15
Gravel	70.	26.66	3.33	5.57	5.57
Dirt pebbles	10.	86.67	3.33	5.75	4.73
Snow	80.	16.67	3.33	5.81	5.67
High Grass	0.	83.33	16.67	-	5.4
Forest Underbrush	63.33	33.33	3.33	4.84	4.7
Dry Leaves	40.	60.	0.	4.75	4.96
Wood	46.67	20.	33.33	3.93	4.55
Creaking Wood	93.33	6.67	0.	5.16	5.17
Metal	80.	13.33	6.67	3.33	4.77
Wood plus Reverb	20.	70.	10.	3.83	4.76
Creaking Wood plus Reverb	93.33	3.33	3.33	4.93	5.17
Metal plus Reverb	83.33	10.	6.67	3.6	5.23

Table 1: Results of experiment 1: recognition (in percentage) of the surfaces with the interactive system.

4.1.3 Task

During experiment 1 the participants were asked to wear a pair of headphones and to walk in the area delimited by the microphones. They were given the list of different surfaces to be held in one hand, presented as non-forced alternate choice.

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

1. Which surface do you think you are walking on? For each stimulus choose an answer in the following list: 1) beach sand, 2) gravel, 3) dirt plus pebbles, 4) snow, 5) high grass, 6) forest underbrush, 7) dry leaves, 8) wood, 9) creaking wood, 10) metal, 11) carpet, 12) concrete, 13) frozen snow, 14) puddles and water, 15) I don't know.

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).

At the end of the experiment, subjects 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.

The task in experiment 2 was similar to experiment 1. However, subjects were sitting on a chair, listening to the sounds through headphones and interacting with the interface mentioned in section 4.1.2. The task in experiment 3 was similar to experiment 2, but in addition to the classification of the surfaces subjects were also asked to evaluate the degree of certainty of their choice on a scale from 1 to 7. At the end of the experiments 2 and 3 the subjects were also given the opportunity to leave an open comment on their experience interacting with the system.
The list included a range of materials wider than those presented in experiment 1 and 2 (see section 4.1). Conversely, in experiment 3 all the materials in the list were presented. The subjects were informed that they could choose the same material more than one time and that they were not forced to choose all the materials in the list. In addition for experiment 1, they could use the interactive system as much as they wanted before giving an answer. Likewise for experiments 2 and 3 they could listen to the sounds as much as they wanted. When passed to the next stimulus they could not change the answer to the previous stimuli.

4.2 Results

The collected answers were analyzed and compared between the three experiments. Results concerning the percentage of correct answers in experiment 1 and 2 are illustrated in Figure 3, while the comparison between the two experiments in terms of realism and quality of the sounds is showed in Figure 4. The degree of realism was calculated only looking at data from correct answers, i.e., when the surface was correctly recognised.

The first noticeable element emerging from both figures is that almost always the use of the interactive system gave rise to a better recognition of the surfaces and a higher evaluation of realism and quality of the proposed sounds, rather than the recorded sounds.
In both experiments the footsteps sounds on snow, creaking wood (with and without reverb), gravel and metal (with reverberation) were correctly recognized with high percentage, while the recognition of the surfaces dirt plus pebbles, high grass and wood (with reverberation) turned out to be wrong most of the times. Regarding the recognition of the other surfaces, good results were found for beach sand and forest underbrush, while correct recognition for dry leaves and wood (without reverberation) were under 50%.

All percentages were higher in experiment 1, although an in-depth analysis shows significant difference only for dry leaves (χ^2 = 4.1761, df = 1, p-value = 0.041) and metal (χ^2 = 4.6886, df = 1, p-value = 0.03036).

An analysis performed on the wrong answers reveals that in average subjects tended to classify erroneously a surface as another belonging to a same category (e.g., wood-concrete, snow-frozen snow, dry leaves-forest underbrush) rather than to different categories (e.g., wood-water, wood-gravel, metal-dry leaves).
Moreover, results show that the addition of the reverberation to the sounds gave rise to better recognitions for metal, and worse for wood plus reverberation, which was perceived most of the times as concrete (not tangible differences were found for the creaking wood). As concerns the comparisons between reverberated and not reverberated sounds in terms of realism negligible differences were found, while in terms of quality the reverberated sounds led to light higher evaluations.

Results of the third experiment are illustrated in Figure 5. Recognition of recorded sounds was quite good in average, with a better performance for the solids and liquids surfaces compared to the aggregate ones. In particular metal, wood, creaking wood, concrete, frozen snow, water and gravel show very high percentages, while the sound of the high grass confirms the negative trend already emerged in the previous experiments. All the other materials present percentages over 50%, with the exception of dry leaves, dirt plus pebbles and forest underbrush, as the data in Table 3 show.

	Correct answers	Wrong answers	I don't know	Realism	Quality
Beach Sand	36.67	63.33	0.	4.18	4.77
Gravel	66.67	33.33	0.	4.9	4.57
Dirt pebbles	10.	90.	0.	4.71	4.37
Snow	83.33	13.33	3.33	5.17	5.14
High Grass	3.33	93.33	3.33	5.	4.3
Forest Underbrush	36.67	60.	3.33	4.22	4.27
Dry Leaves	13.33	86.67	0.	4.5	3.87
Wood	36.67	33.33	30.	4.58	4.
Creaking Wood	76.67	13.33	10.	3.17	3.82
Metal	50.	16.67	33.33	2.93	3.29
Wood plus Reverb	6.67	73.33	20.	6.	4.04
Creaking Wood plus Reverb	76.67	13.33	10.	3.35	3.89
Metal plus Reverb	66.67	6.67	26.67	3.3	4.3

Table 2: Results of experiment 2: recognition (in percentage) of the surfaces with the recorded synthesized sounds.

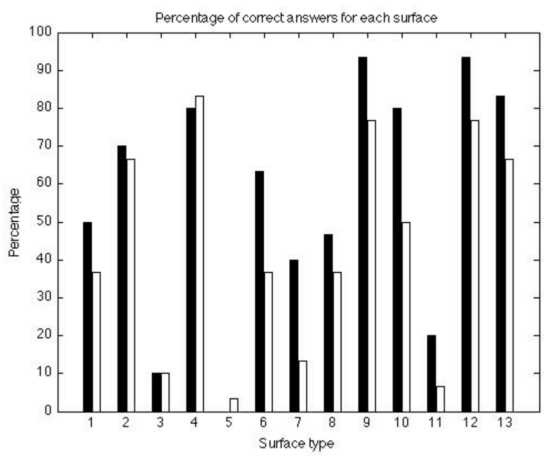

Figure 3: Comparison of the percentages of correct answers for each surface in experiment 1 (black) and 2 (white). Surface type from left to right: 1-beach sand, 2-gravel, 3-dirt pebbles, 4-snow, 5-high grass, 6-forest underbrush, 7-dry leaves, 8-wood, 9-creaking wood, 10-metal, 11-wood plus reverberation, 12-creaking wood plus reverberation, 13-metal plus reverberation. Notice that the missing element in column 5 indicates the fact that none of the subjects was able to recognize the high-grass in experiment 1.

The degree of certainty in the answers seems to be on average consistent with the percentage of correctness (even if there are some exceptions, as for the footsteps on puddles, which were erroneously classified as footsteps through the water).

What emerges from these results is the ability of the subjects in distinguishing materials in the same category for solid surfaces, and their difficulties in the recognition of aggregate surfaces (aspect also confirmed by the comments of the participants). Indeed the analysis of the wrong answers for aggregate surfaces confirms the tendency already showed in the previous experiments, in classifying erroneously a surface as another belonging to a same category.

From the comparison with the results of the recognition of the surfaces presented in the previous two experiments, one can note that the percentage of correct answers for the same surfaces is higher for the experiment 3, with the exception of the snow in both experiment 1 and 2, and of forest underbrush and creaking wood (with reverberation) in experiment 1. Moreover similar percentages were found for beach sand in experiment 1 and 3, and the percentage of gravel is high in the same experiments. Finally the very low percentages for the high grass in the three experiments confirm that this is a sound difficult to recognize.

The final questions of experiment 1 (evaluated on a seven-point Likert scale) show that subjects judged the interaction with the system quite natural (mean= 5.6), and that they felt quite normal (mean= 5.33) and a little bit constrained (mean 2.9) during the act of walking. Indeed, subjects commented on the need of a wider area to walk and of a wireless headphones set.

Finally, regarding the "I do not know" answers the percentage was higher in experiment 2 (10.77%) rather than experiment 1 (7.95%), and lower (3.11%) for the experiment 3. Tables 1,2, 3 show in details the results of experiment 1, 2 and 3 respectively.

5 Discussions

A footstep sound is extremely dependent both on the kind of shoes a person is wearing and on the kind of floor the person is walking on. All sounds were synthesized assuming that the shoes hitting the floor had a solid sole. This aspect is extremely important in the simulation of solid floors. As a matter of fact, when interacting with virtual wood and metal more than one participant commented of having the sensation of wearing a different kind of shoe. More precisely, they commented that they felt like they were wearing a shoe with a solid sole. This indicates the ability of auditory feedback to affect perception of material.

homogeneous ones: as proof of the good success our design we

	Correct answers	Wrong answers	I don't know	Degree of certainty
Beach Sand	53.34	46.67	0.	4.5
Gravel	86.67	13.33	0.	5.07
Dirt pebbles	43.33	56.67	0.	4.8
Snow	63.33	36.67	0.	5.93
Frozen Snow	70.	16.67	13.33	5.66
High Grass	10.	83.33	6.67	3.5
Forest Underbrush	36.67	60.	3.33	4.14
Dry Leaves	56.67	43.33	0.	4.77
Concrete	70.	23.33	6.67	4.43
Wood	80.	16.67	3.33	5.17
Creaking Wood	90.	6.67	3.33	6.14
Metal	96.67	3.33	0.	6.33
Carpet	66.67	23.33	10.	4.78
Puddles	33.33	66.67	0.	5.8
Water	86.67	13.33	0.	6.57

Table 3: Results of experiment 3: recognition (in percentage) of the surfaces with the recorded real sounds.

found that for the sound of the wood and metal floors more than one participant commented that he/she felt like wearing a different kind of shoe, and for the precision with a solid soil.

In general, the use of the interactive system facilitated the recognition task, and the sound quality of the system was perceived as higher.

One peculiar element of the interactive system is the lack of haptic feedback which is present when walking in the real world and is an important element in the perception of a surface. This lack will be compensated in future implementations of the system, where haptic feedback will be integrated. Some subjects also commented on the importance of visual feedback, which would have obviously helped in the recognition task.

6 Conclusions and future work

In this paper, we introduced a real-time footsteps synthesizer controlled by the user, which works independently from the footwear. This is a feature that distinguishes our prototype from other systems developed with similar goals.

The system was tested in a between-subjects experiment, where it was compared to a recognition task including recorded and synthesized offline sounds. Results show that subjects are able to recognize most of the synthesized surfaces using the interactive system with high accuracy. Similar accuracy can be noticed in the recognition of real recorded footsteps sounds, which is an indication of the success of the proposed algorithms and their control.

The developed system is ready to be integrated in computer games and interactive installations where a user can navigate. The simulations proposed, however, reproduce the act of walking on a flat surface.

On the other hand, real life scenarios include also uphill and downhill movements whose footsteps sounds differ significantly from those produced while walking on a flat surface. Such situations can be incorporated in our synthesis engine, by modifying different parameters of the corresponding sounds such as amplitude and temporal variations.

In future work, we indeed plan to utilize the system in multimodal environments, and include haptic and 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.

7 Acknowledgments

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. [6]

References

[1] J. Adrien. *The missing link: Modal synthesis*. MIT Press Cambridge, MA, USA, 1991.

[2] F. Avanzini and D. Rocchesso. Modeling collision sounds: Non-linear contact force. In *Proc. COST-G6 Conf. Digital Audio Effects (DAFx-01)*, pages 61–66, 2001.

[3] A. Benbasat, S. Morris, and J. Paradiso. A wireless modular sensor architecture and its application in on-shoe gait analysis. In *Sensors, 2003. Proceedings of IEEE*, volume 2, 2003.

[4] P. Chueng and P. Marsden. Designing Auditory Spaces to Support Sense of Place: The Role of Expectation. In *CSCW Workshop: The Role of Place in Shaping Virtual Community*. Citeseer, 2002.

[5] P. Cook. Physically Informed Sonic Modeling (PhISM): Synthesis of Percussive Sounds. *Computer Music Journal*, 21(3):38–49, 1997.

[6] P. Cook. Modeling Bill's Gait: Analysis and Parametric Synthesis of Walking Sounds. *Proceedings of the AES 22nd International Conference on Virtual, Synthetic, and Entertainment Audio*, pages 73–78, 2002.

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

[8] P. Dupont, V. Hayward, B. Armstrong, and F. Altpeter. Single state elastoplastic friction models. *IEEE Transactions on Automatic Control*, 47(5):787–792, 2002.

[9] A. Farnell. Marching onwards: procedural synthetic footsteps for video games and animation . *Proceedings of the Pure Data Convention*, 2007.

[10] F. Fontana and R. Bresin. Physics-based sound synthesis and control: crushing, walking and running by crumpling sounds. *Proc. Colloquium on Musical Informatics*, pages 109–114, 2003.

[11] J. Freeman and J. Lessiter. Hear there & everywhere: the effects of multi-channel audio on presence. *Proceedings of ICAD 2001*, pages 231–234, 2001.

[12] W. Gaver. What in the world do we hear?: An ecological approach to auditory event perception. *Ecological Psychology*, 5(1):1–29, 1993.

[13] B. Giordano and S. Mcadams. Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates. *The Journal of the Acoustical Society of America*, 119:1171, 2006.

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

[6] www.niwproject.eu

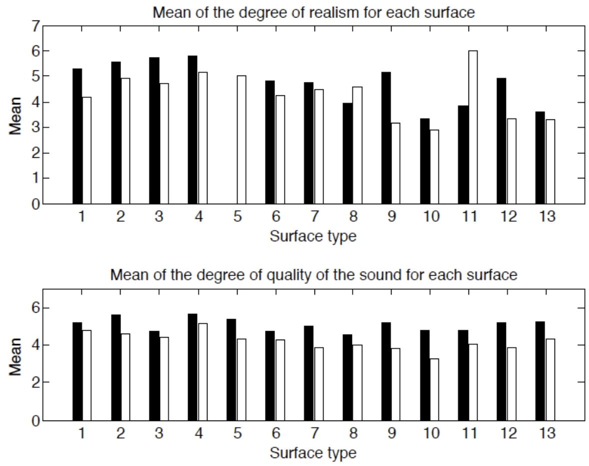

Figure 4: Comparison of the mean of the degree of realism (top) and quality of the sound (bottom) for each surface in experiment 1 (black) and 2 (white). Surface type from left to right: 1-beach sand, 2-gravel, 3-dirt pebbles, 4-snow, 5-high grass, 6-forest underbrush, 7-dry leaves, 8-wood, 9-creaking wood, 10-metal, 11-wood plus reverberation, 12-creaking wood plus reverberation, 13-metal plus reverberation.

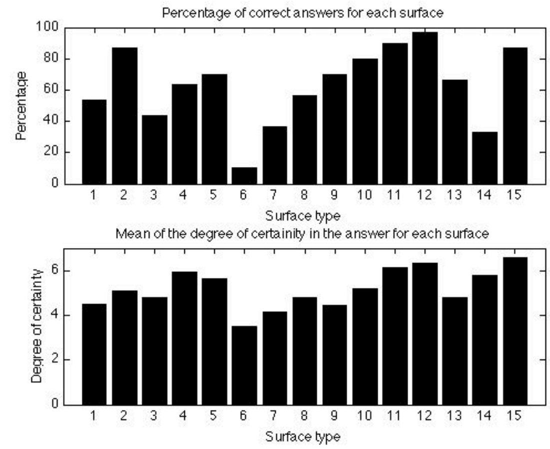

Figure 5: Percentages of correct answers (top) and mean of the degree of certainity in the answer for each surface in experiment 3. Surface type from left to right: 1-beach sand, 2-gravel, 3-dirt pebbles, 4-snow, 5-frozen snow, 6-high grass, 7-forest underbrush, 8-dry leaves, 9-concrete, 10-wood, 11-creaking wood, 12-metal, 13-carpet, 14-puddles and 15-water.

[15] B. Kapralos, D. Zikovitz, M. Jenkin, and L. Harris. Auditory cues in the perception of self-motion. In *116th AES convention*. Citeseer, 2004.

[16] R. Klatzky, D. Pai, and E. Krotkov. Perception of material from contact sounds. *Presence: Teleoperators & Virtual Environments*, 9(4):399–410, 2000.

[17] P. Larsson, D. Västfjäll, and M. Kleiner. Perception of self-motion and presence in auditory virtual environments. In *Proceedings of seventh annual workshop presence*, pages 252–258, 2004.

[18] A. Law, B. Peck, Y. Visell, P. Kry, and J. Cooperstock. A Multi-modal Floor-space for Experiencing Material Deformation Underfoot in Virtual Reality. In *IEEE International Workshop on Haptic Audio visual Environments and Games, 2008. HAVE 2008*, pages 126–131, 2008.

[19] X. Li, R. Logan, and R. Pastore. Perception of acoustic source characteristics: Walking sounds. *The Journal of the Acoustical Society of America*, 90:3036, 1991.

[20] R. Nordahl. Increasing the motion of users in photorealistic virtual environments by utilizing auditory rendering of the environment and ego-motion. *Proceedings of Presence*, pages 57–62, 2006.

[21] J. Paradiso, K. Hsiao, and E. Hu. Interactive music for instrumented dancing shoes. In *Proc. of the 1999 International Computer Music Conference*, pages 453–456, 1999.

[22] R. Pastore, J. Flint, J. Gaston, and M. Solomon. Auditory event perception: The source-perception loop for posture in human gait. *Perception and Psychophysics*, 70(1):13, 2008.

[23] N. Raghuvanshi and M. Lin. Interactive sound synthesis for large scale environments. In *Proceedings of the 2006 symposium on Interactive 3D graphics and games*, pages 101–108. ACM New York, NY, USA, 2006.

[24] R. Sanders Jr. *The effect of sound delivery methods on a users sense of presence in a virtual environment*. PhD thesis, NAVAL POSTGRADUATE SCHOOL, 2002.

[25] R. Storms and M. Zyda. Interactions in perceived quality of auditory-visual displays. *Presence: Teleoperators & Virtual Environments*, 9(6):557–580, 2000.

[26] A. Väljamäe, P. Larsson, D. Västfjäll, and M. Kleiner. Travelling without moving: Auditory scene cues for translational self-motion. In *Proceedings of ICAD05*, 2005.

[27] K. Van Den Doel, P. Kry, and D. Pai. FoleyAutomatic: physically-based sound effects for interactive simulation and animation. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 537–544. ACM New York, NY, USA, 2001.

[28] D. Västfjäll. The subjective sense of presence, emotion recognition, and experienced emotions in auditory virtual environments. *CyberPsychology & Behavior*, 6(2):181–188, 2003.