An Investigation on Temporal Aspects in the Audio-Haptic Simulation of Footsteps

Luca Turchet and Stefania Serafin

Department of Architecture, Design and Media Technology,
Aalborg University Copenhagen
{tur,sts}@create.aau.dk

Abstract. In this paper, we present an experiment whose goal is to assess the role of temporal aspects in sonically and haptically simulating the act of walking on a bump or a hole. In particular, we investigated whether the timing between heel and toe and the timing between footsteps affected perception of walking on unflat surfaces. Results show that it is possible to sonically and haptically simulate a bump or a hole only by varying temporal information.

Keywords: Footstep sounds, physical models, auditory feedback, haptic feedback.

1 Introduction

Previous research on simulating walking sounds using physics based engines has focused on the act of walking on flat surfaces [2,4,3,14,11]. In the virtual reality community, few locomotion interfaces are able to render uneven grounds, and they have the disadvantage of being costly and cumbersome [5,6,8]. Recently, research has shown that it is possible to simulate the act of walking on unflat surfaces by only using visual cues [10]. Three parameters of camera motion were considered in the simulation: orientation, velocity and height, and their combination. The experiments were run both actively, having users wear an head mounted display, as well as passively, having users look at a video of the simulations. Results show that such visualization techniques successfully simulate bumps and holes located in the ground. These results are a development of previous research on pseudo-haptic simulation [9]. This research was extended by implementing a multimodal (audio-visual) simulation of walking on a bump or a hole [16]. Results in this case showed that the auditory cues reinforce the visual cues when coherent cues are provided in both modalities. When subjects were exposed to conflicting cues, for example by simulating visually the act of walking on a bump and auditorily the act of walking on a hole, usually the visual cues are dominant, apart from when the velocity effect is the visual parameter varied. This might be due to the higher temporal resolution of the auditory system versus the visual system [21].

In this paper, we are interested in exploring the possibility of implementing such pseudo-haptic feedback from the sonic and haptic point of view. Recently,

A. Vatakis et al. (Eds.): Time and Time Perception 2010, LNAI 6789, pp. 101–115, 2011.

we developed a system which can provide combined auditory and haptic sensations that arise while walking on aggregate and solid surfaces. The system is composed of an audio-haptic synthesis engine, and a pair of shoes enhanced with sensors and actuators. Such engine is based on physical models, that drive both the haptic and audio synthesis. A complete description of such system and of all its components is given elsewhere in detail [12], [17].

In a previous study [15], we used the synthesis engine in order to run an experiment whose goal was to assess the role of temporal aspects in sonically simulating the act of walking on a bump or a hole. In particular, we investigated whether the timing between heel and toe and the timing between footsteps affected perception of walking on unflat surfaces. In the experiment the footstep sounds where prerecorded. Results showed that it is possible to simulate a bump or a hole only by using temporal information [15].

Starting from those results, in this paper we are interested in understanding whether at haptic level it is possible to simulate the act of walking on a bump or a hole. Such haptic information is generated by means of the same techniques applied in [15] for the auditory simulation of bumps and holes, i.e., varying the temporal distances between the vibrations corresponding to the heel and to the toe as well as those between footsteps.

2 Simulation Hardware and Software

In this section we briefly describe the system used in the experiments presented in this paper. As mentioned in section 1, the complete description of such system can be found in our previous research [20], [17].

We developed a system which simulates both offline and in real-time the auditory and haptic sensation of walking on different surfaces. Specifically, the sensation of walking on solid surfaces is simulated by using and impact model [7], while to simulate walking on aggregate grounds, we used a physically informed sonic models (PhiSM) algorithm [1].

In order to provide both audio and haptic feedback, haptic shoes enhanced with pressure sensors have been developed. In detail, such shoes are pair of light-weight sandals (Model Arpenaz-50, Decathlon, Villeneuve d'Ascq, France). This particular model has light, stiff foam soles that are easy to gouge and fashion. Four cavities were made in the tickness of the sole to accommodate four vibrotactile actuators (Haptuator, Tactile Labs Inc., Deux-Montagnes, Qc, Canada). These electromagnetic recoil-type actuators have an operational, linear bandwidth of 50–500 Hz and can provide up to 3 G of acceleration when connected to light loads. As indicated in Fig. 1, two actuators were placed under the heel of the wearer and the other two under the ball of the foot. They were bonded in place to ensure good transmission of the vibrations inside the soles. When activated, vibrations propagated far in the light, stiff foam. The sole has force sensors intended to pick the foot-floor interaction force in order the drive the audio and haptic synthesis. They were not used in the present study.

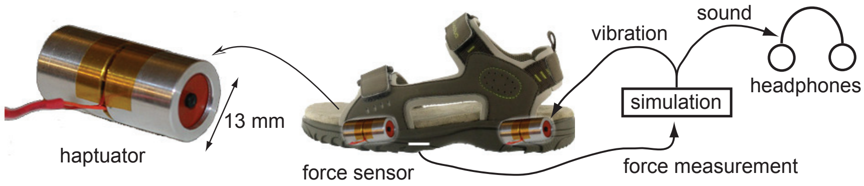

Fig. 1. System (one shoe shown). Left: recoil-type actuation from Tactile Labs Inc. The moving parts are protected by an aluminum enclosure able to bear the weight of a person. Middle: approximate location of the actuators in the sandal. Right: system diagram showing the interconnections. Here the force signal was not used.

The involved hardware provides a control in real-time of an audio-haptic synthesis engine. Such engine is based on physical models which are driven by a signal, in the audio domain, expressing the ground reaction force (GRF), i.e., the reaction force supplied by the ground at every step. In our simulations the GRF corresponds to the amplitude envelope extracted from an audio signal containing a footstep sound.

As previously mentioned, the engine can work both offline and in realtime. The two approaches differ for the way the input GRF is generated. Concerning the realtime implementation, various systems for the generation of such input have been developed and tested [20], [17], [12], [13], [18], [19]. In the offline implementation, the input signal consists of an audio file from which the GRF is extracted. Such file consists of a recording of a person walking on a real surface. Better results in terms of the GRF detection can be found on audio recordings of walking on solid surfaces and with a small amount of background noise. The different envelope profiles of each step in the file are extracted and fed to the engine which produces the synthesized footstep sounds according to the choice of the surface to be simulated.

In the experiment presented in this paper, we adopted the offline use of the engine. To control the engine, we created different audio files placing at various temporal patterns the recording of an unique real footstep sound on concrete. Such sound was chosen among those available in the Hollywood Edge sound effects library.[1] Three types of surface profiles have been created starting from the footstep sound generator: bumps, holes, and flat surfaces (see Figure 2). The details about the simulation of such surface profiles are shown in section 3.4. For the purpose of the experiment, two types of surfaces, one solid (wood) and one aggregate (gravel), were chosen. The reason for choosing two materials was to assess whether the surface type affected the quality of the results.

3 Experiment Design

We conducted one experiment whose goal was to ascertain the participants' ability to recognize if the sounds and the haptic sensations they were exposed to

[1] Hollywood Edge sound effects library: `www.hollywoodedge.com`

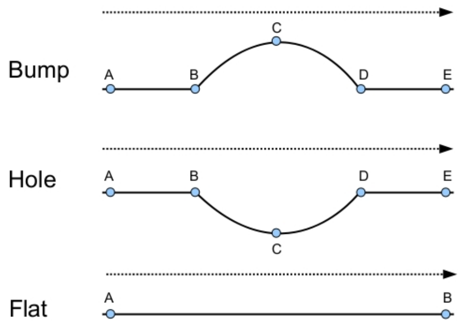

Fig. 2. The three types of surfaces modeled

corresponded to walking on a bump, a hole or a flat surface. Specifically, subjects were exposed to one of the three following conditions:

1. Condition 1: auditory simulation of bumps and holes
2. Condition 2: haptic simulation of bumps and holes
3. Condition 3: audio-haptic simulation of bumps and holes

The results of the experiment related to the first condition have been already published in our previous research [15] and they are illustrated in section 4.1. In this paper we are interested in extending to the haptic level those results achieved with the auditory cues. Indeed the technique used for the simulation of the three types of surface profiles at haptic level is exactly the same proposed for the auditory level. In detail, the audio files used as input to the audio-haptic synthesis engine were created placing at different temporal intervals the footstep sound generator (see section 3.4). The basic idea is that in a real environment, a person generally walks slower on ascending slopes, and faster on descending slopes. We transposed this information in our experiment by modifying the time intervals both between footsteps and between the heel and toe information in each footstep.

The goal of condition 2 was to assess whether at haptic level it is possible to simulate the act of walking on a bump or a hole by means of this technique, when subjects are not walking but are sitting on a chair and passively receive the haptic stimuli. In this haptic condition, a noise signal masked the auditory input generated by the haptic shoes as result of the activation of the actuators. In particular, participants were presented with a continuous 70 dB spl pink noise over the headphones described in section 3.2. In condition 3, we were interested in understanding whether the addition of haptic feedback to the auditory one enhanced the recognition of the proposed surface profiles as well as the realism of the simulation.

3.1 Participants

Forty-five participants were divided in three groups (n=15) to perform the experiment. The three groups were composed respectively of 11 men and 4 women, aged between 20 and 29 (mean=23.6, standard deviation=2.84), 11 men and 4 women, aged between 21 and 32 (mean=24.86, standard deviation=3.48) and 11 men and 4 women, aged between 20 and 28 (mean=23.06, standard deviation=2.40). All participants reported normal hearing conditions. All participants were naive with respect to the experimental setup and to the purpose of the experiment. The participants took on average about 15, 17 and 21 minutes for experiments 1, 2 and 3 respectively.

3.2 Setup

The experiment was carried out in an acoustically isolated laboratory where the setups for the experiments were installed. They consisted of a simple graphical user interface with which participants were asked to interact, and a spreadsheet to collect their answers. The interface was created using the Max/MSP program[2] and was composed only by buttons to be pressed. Each button was numbered, and by pressing it an audio or audio-haptic stimulus was triggered and conveyed to the user by means of headphones[3] and the haptic shoes respectively. The choice of delivering auditory feedback using headphones was motivated by the fact that we wanted the subjects to be isolated from external noise. Users were asked to press each button according to their numerical order, and to write the corresponding answers on the spreadsheet.

3.3 Task

During the experiment subjects were sitting on a chair, listening to the sounds through headphones and feeling the haptic vibrations through the haptic shoes, and interacting with the interface mentioned in section 3.2.

They were given the list of three different surfaces (bump, hole, flat), presented as forced alternate choice. The task consisted of recognizing to which surface the walk corresponded after the presentation of the stimulus. 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 (1=very low certainty, 7=very high certainty). Furthermore, in conditions 2 and 3 they were asked to evaluate, again on a seven point Likert scale, the degree of realism and the degree of quality of the perceived stimulus. Participants were allowed to listen to the sounds as much as they wanted before giving an answer. When moving to the next stimulus they could not change the answer to the previous stimuli.

3.4 Haptic and Auditory Simulation

As described in section 3, the technique used for the simulation of bumps, holes and flat surface consisted of temporal intervals variations. The temporal patterns used were designed to simulate 14 different surface profiles. Specifically 2 flat, 6

[2] Max/MSP: `www.cycling74.com`
[3] Sennheiser HD 600, `http://www.sennheiser.com`

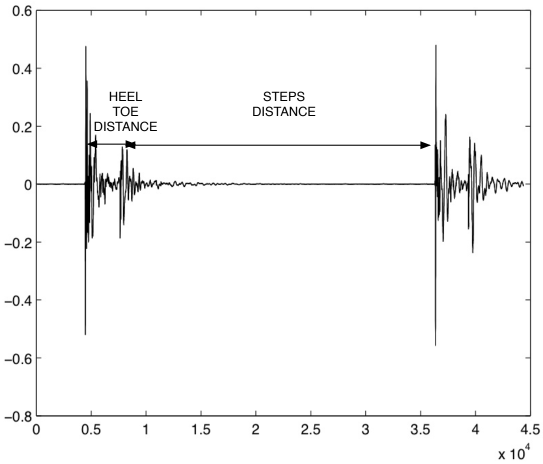

Fig. 3. Temporal distances between (named "steps distance" in the Figure), and within (named "heel-toe distance" in the Figure) footsteps. x-axis: Time (s). y-axis: Amplitude.

bumps and 6 holes were designed. Such patterns involved three types of temporal distances. The first was the temporal distance between footsteps (i.e., the time interval between the end of the sound generated by the toe and the beginning of the sound generated by the heel of the next step), the second was the temporal distance between heel and toe (i.e., the time interval between the end of the sound generated by the heel and the beginning of the sound generated by the toe in the same step), the third consisted of the combination of the previous two (see Figure 3).

The characteristics of the 14 files used to drive the sound engine are illustrated in table 1. In such table the suffixes _step, _h_t and _comb indicate the type of temporal distance used for each file (footsteps distance, heel-toe distance and their combinations respectively). The equations in the "Number of steps" column indicate how the steps where placed in reference to Figure 2. As an example, the stimulus bump_2_step was composed by 19 steps, 4 steps to go from point A to point B, 6 steps to go from point B to point C, 5 steps to go from point C to point D, and 4 steps to go from point D to point E). In order to model two different types of bumps and holes, for each category of surface modeling (by means of the three temporal distance types), two slopes where chosen. In all the three conditions participants were exposed to 28 trials, where 14 surface profiles were presented twice in randomized order. In condition 1 and 2 (see sections 4.1 and 4.2) the simulated surface profiles are all those presented in table 1, while in condition 3 only 7 of them was used and presented both in audio and audio-haptic conditions (see section 4.3).

The audio-haptic synthesis engine was set in order to synthesize footstep sounds and vibrations on two different kinds of materials: wood and gravel. Each surface profile was presented with both wood and gravel.

Table 1. Features of the 14 files used as input to the audio-haptic engine. For a detailed description, see the text.

	Duration (sec.)	**Number of steps**	**Footsteps temporal distance increment (ms.)**	**Footsteps temporal distance range (ms.)**	**Heel-toe increment (ms.)**	**Heel-toe range (ms.)**
flat_1	12	19	-	550 (fixed)	-	69 (fixed)
flat_2	16	19	-	750 (fixed)	-	69 (fixed)
bump_1_step	27	31 = 4+12+11+4	50	550 ⟶ 1150	-	69 (fixed)
bump_2_step	16	19 = 4+6+5+4	100	550 ⟶ 1150	-	69 (fixed)
hole_1_step	18	31 = 4+12+11+4	-50	750 ⟶ 150	-	69 (fixed)
hole_2_step	11	19 = 4+6+5+4	-100	750 ⟶ 150	-	69 (fixed)
bump_1_h_t	24	31 = 4+12+11+4	-	550 (fixed)	30	0 ⟶ 360 (+ 69)
bump_2_h_t	14	19 = 4+6+5+4	-	550 (fixed)	60	0 ⟶ 360 (+ 69)
hole_1_h_t	24	31 = 4+12+11+4	-	550 (fixed)	-20	240 ⟶ 0 (+ 69)
hole_2_h_t	13	19 = 4+6+5+4	-	550 (fixed)	-40	240 ⟶ 0 (+ 69)
bump_1_comb	32	31 = 4+12+11+4	50	550 ⟶ 1150	30	0 ⟶ 360 (+ 69)
bump_2_comb	19	19 = 4+6+5+4	100	550 ⟶ 1150	60	0 ⟶ 360 (+ 69)
hole_1_comb	22	31 = 4+12+11+4	-50	750 ⟶ 150	-20	240 ⟶ 0 (+ 69)
hole_2_comb	15	19 = 4+6+5+4	-100	750 ⟶ 150	-40	240 ⟶ 0 (+ 69)

The haptic signal was separated in order to activate coherently the actuators placed under the heel and the toe. In our simulation the actuators placed under the heel of the right foot were activated simultaneously with those of the left foot, using the same signal. Analogously, the haptic signal for the toes was conveyed simultaneously to the actuators placed under the toes of both the feet.

4 Results

In this section we show the results for condition 1 (audio condition), condition 2 (haptic condition) and condition 3 (comparison between audio and audio-haptic conditions). In section 4.1 we report the results found in a previous work [15], where the 14 surface profiles described in section 3.4 were used to simulate a bump or a hole only by using auditory cues.

4.1 Results of Condition 1 (Audio Condition)

The results of the condition 1 for wood and gravel are shown in tables 2 and 3 respectively. In both tables, the first column shows the different conditions as described in table 1. The second, third and fourth columns illustrate the choices of the subjects (bump, hole or flat) for the different conditions they were exposed to. The fifth, sixth and seventh column report the average certainty expressed by the subjects after performing their choice; the fifth column reports the total certaintly in both correct and wrong answers, while the sixth and seventh column report the certainty in correct and uncorrect answers respectively. Finally, the last column reports the percentage of correct answers.

As the tables show, subjects could successfully recognize bumps and holes using only the auditory cues described in the previous section. In fact, as can be seen in the last column of tables 2 and 3, the percentage of correct answers is

Table 2. Results of condition 1 (audio condition) for the wood surface

	Bump	Hole	Flat	Mean certainty Total	Mean certainty Correct answers	Mean certainty certainty Wrong answers	% Correct answers
flat_1	1		14	6	6.1429	4	93.33
flat_2		1	14	5.6667	6	1	93.33
bump_1_step	13	2		5.4286	5.5385	4.5	93.33
bump_2_step	14	1		5.4667	5.5714	4	93.33
hole_1_step	1	14		4.8	4.7857	5	93.33
hole_2_step	2	13		4.9333	5	4.5	86.66
bump_1_h_t	13	1	1	5.6	5.8469	4	86.66
bump_2_h_t	12	2	1	5.1333	5.4167	4	80
hole_1_h_t		15		4.2667	4.2667		100
hole_2_h_t	3	12		4.4	4.8182	3.25	80
bump_1_comb	14	1		5.4	5.5	4	93.33
bump_2_comb	14	1		5.1333	5.3571	2	93.33
hole_1_comb	1	14		5.1333	5.2143	4	93.33
hole_2_comb	1	14		4.9333	5	4	93.33

Table 3. Results of condition 1 (audio condition) for the gravel surface

	Bump	Hole	Flat	Mean certainty Total	Mean certainty Correct answers	Mean certainty Wrong answers	% Correct answers
flat_1		3	12	5.2667	5.6667	3.6667	80
flat_2	1	1	13	5.4667	6	2	86.66
bump_1_step	12	2	1	5.8469	5.75	4.6667	80
bump_2_step	13	1	1	5.4667	5.9231	2.5	86.66
hole_1_step		14	1	5.4667	5.7857	1	93.33
hole_2_step		15		6	6		100
bump_1_h_t	14	1		5.5333	5.7857	2	93.33
bump_2_h_t	11	2	2	4.5333	4.8182	3.75	73.33
hole_1_h_t	1	12	2	4.5333	4.5833	4.3333	80
hole_2_h_t		15		4.4	4.4		100
bump_1_comb	13	2		5.3333	5.6667	4	86.66
bump_2_comb	13	1	1	5.2667	5.7692	2	86.66
hole_1_comb	2	12	1	5.2667	5.3333	5	80
hole_2_comb	2	13		5.2667	5.8469	1.5	86.66

high for all conditions, reaching also 100 % of correct answers in three conditions, and with a lowest score of 73 % which was reached only in one condition.

Observing columns 6 and 7, morever, it is possible to notice how subjects are quite certain when they express a correct answer. In both surfaces, indeed, the mean certainty for correct answers is always above average. On the other hand, in situations where the answer was incorrect the degree of certainty is also extremely low. This is the case, for example, in the second flat stimulus for the wood surface and the first hole stimulus in the gravel surface.

A t-test was performed to examine whether significant differences were present in the recognition rate among the two surfaces and among the different conditions in the same surface. Overall, no significant differences were measured in the recognition rate among the two surfaces. Moreover, no significant differences were measured in the recognition rate for the different conditions in the same material. For example, no difference was measured in the recognition rate of the first simulated bump footstep versus the second simulated bump footstep. No significant difference was furthermore measured between the recognition rate obtained when changing the temporal information between footsteps versus the one obtained when changing the temporal information within footsteps. Also, the combination of the two temporal information did not significantly enhance the recognition of a bump or a hole. This, however, is also due to the fact that the temporal information taken individually already provided a high recognition rate.

4.2 Results of Condition 2 (Haptic Condition)

Tables 4 and 5 illustrate the results of condition 2 for wood and gravel respectively. The structure of these tables is the same of tables 2 and 3, with the addition of the two columns expressing the average scores of the degree of realism and of the degree of quality.

The first noticeable element emerging from both tables is that it is possible to simulate bumps, holes and flat surfaces by means of haptic stimuli using only the variation of the temporal information. Indeed, on average, subjects could successfully recognize the simulated surfaces using only the haptic cues described in section 3.4, both with wood and gravel. As a matter of fact, the percentage of correct answers is quite high in most of the conditions, reaching also 100 % of correct answers in two conditions, while only in three conditions the score is under 50 %.

From a comparison between all the conditions it is possible to notice that on average the stimuli involving only the variations in the heel-toe distance (i.e., those with suffix _h_t) are recognized with the lowest scores in comparison with the other conditions. Such tendency is also confirmed by values in column 5 which expresses the total degree of certainty for each stimulus (both of correct and wrong answers) and therefore it is an indication of the difficulty of the choice. Indeed in both the tables such values are on average lower for the stimuli with suffix _h_t than all the other conditions.

A chi-square test and a t-test were performed to examine whether significant differences were present in the recognition rate among the two surfaces and among the different conditions in the same surface. Overall, no significant differences were measured in the recognition rate among the two surfaces, nor in the recognition rate for the different conditions in the same material. Moreover no significant difference was measured between the recognition rate obtained when changing the temporal information between footsteps versus the one obtained when changing the temporal information within footsteps. In addition, the combination of the two temporal information did not significantly enhance the recognition of a bump or a hole. Finally, a t-test performed on the realism and on the quality evaluations, did not reveal any significant difference between the two materials nor among the conditions.

Table 4. Results of condition 2 (haptic condition) for the wood surface

	Bump	Hole	Flat	Mean certainty Total	Mean certainty Correct answers	Mean certainty Wrong answers	% Correct answers	Mean Realism	Mean Quality
flat_1	2		13	5.2	5.3077	4.5	86.66	3.9231	4
flat_2	2		13	5.1333	5.3077	4	86.66	3.6154	4.1333
bump_1_step	13		2	4.8	4.7692	5	86.66	4.0769	4.0667
bump_2_step	15			5.2	5.2		100	3.9333	3.8
hole_1_step	2	11	2	4.6667	4.4	5.2	73.33	3.0909	3.6
hole_2_step	2	13		4.6667	4.6923	4.5	86.66	3.7692	3.9333
bump_1_h_t	7	5	3	3.6667	4.1429	3.25	46.66	3	3.4
bump_2_h_t	9	6		4.4667	4.7778	4	60	3.4444	3.4667
hole_1_h_t	1	8	6	3.8	4.375	3.1429	53.33	3.25	3.2
hole_2_h_t		12	3	3.8667	4.25	2.3333	80	3.75	3.2
bump_1_comb	11	4		4.2667	4.2727	4.25	73.33	3.1818	3.2
bump_2_comb	11	3	1	4.7333	5.1818	3.5	73.33	4.0909	3.7333
hole_1_comb	2	13		4.4	4.3846	4.5	86.66	3.5385	3.8
hole_2_comb	1	12	2	4.6667	5.3333	2	80	3.25	3.2

Table 5. Results of condition 2 (haptic condition) for the gravel surface

	Bump	Hole	Flat	Mean certainty Total	Mean certainty Correct answers	Mean certainty Wrong answers	% Correct answers	Mean Realism	Mean Quality
flat_1			15	4.4667	4.4667		100	3.6	3.8667
flat_2	1	1	13	5.2667	5.6923	2.5	86.66	4.7692	4.3333
bump_1_step	12		3	4.9333	5	4.6667	80	3.9231	3.7333
bump_2_step	11	2	2	4.4667	4.7273	3.75	73.33	3.3636	3.33
hole_1_step		12	3	4.8	4.9167	4.3333	80	3.6667	3.7333
hole_2_step		12	3	5.5333	5.75	4.6667	80	4.25	4.2
bump_1_h_t	8	6	1	4.4	5	3.7143	53.33	2.875	3.5333
bump_2_h_t	11	1	3	3.8667	4.1818	3	73.33	3.7273	3.8667
hole_1_h_t	1	7	7	4.3333	4.2857	4.375	46.66	3.8571	4
hole_2_h_t	3	6	6	3.7333	3.6667	3.7778	40	3.5	3.4667
bump_1_comb	13	1	1	5.2	5.2308	5	86.66	3.7692	4
bump_2_comb	10	4	1	4.4	5.1	3	66.66	3.5	3.5333
hole_1_comb	1	13	1	4.8	4.8462	4.5	86.66	3.3077	3.4667
hole_2_comb		14	1	4.8	4.9286	3	93.33	4.2143	4

4.3 Results of Condition 3 (Audio and Audio-Haptic Condition)

In this experiment we used only a subset of the 14 files used in the previous conditions. This was due in order to avoid that the experiment became too long for the participants, and therefore started answering randomly. In particular, we selected 7 stimuli to be presented both with the modality audio and audio-haptic (suffixes _A and _AH respectively in tables 6 and 7), and both with wood and gravel. Precisely we used those with suffix _2 in table 1, with the exception of the flat stimulus for which we used the one with suffix _1.

The first noticeable element emerging from both tables is that subjects could successfully recognize the simulated surfaces with high precision both at audio and audio-haptic level. An exception is the stimulus hole_h_t_A with wood material for which the score is 40 %, differing from the percentage obtained by the same stimulus in condition 1. Anyways a chi-square test revealed that such difference is not statistically significant so we can state that all the results concerning the audio modality are consistent with those found in condition 1.

Table 6. Results of condition 3 (audio and audio-haptic condition) for the wood surface

	Bump	Hole	Flat	Mean certainty Total	Mean certainty Correct answers	Mean certainty Wrong answers	% Correct answers	Mean Realism	Mean Quality
flat_A			15	6	6	-	100	3.7333	3.6
flat_AH		1	14	6.0667	6.4286	1	93.33	4.9286	4.4667
bump_step_A	11	2	2	4.7333	5.1818	3.5	73.33	3.2727	3.6667
bump_step_AH	12	2	1	5.2667	5.5833	4	80	3.9167	4.6
hole_step_A	1	14		5.6667	5.7143	5	93.33	3.6429	4
hole_step_AH	2	12	1	5.1333	5.6667	3	80	4.75	4.4667
bump_h_t_A	13	2		4.9333	5.3846	2	86.66	3.6154	3.33
bump_h_t_AH	13	2		4.6667	5.2308	1	86.66	4.3846	4.0667
hole_h_t_A	2	6	7	3.7333	4.5	3.2222	40	2.8571	3
hole_h_t_AH	2	11	2	4.0667	4.1818	3.75	73.33	3.6364	4.4
bump_comb_A	12	3		4.2	4.8333	1.6667	80	3.25	3.6667
bump_comb_AH	14	1		5.0667	5.2857	2	93.33	4.2143	4.3333
hole_comb_A	1	12	2	4.9333	5.25	3.6667	80	3.5	3.8
hole_comb_AH	2	13		5.3333	5.5385	4	86.66	4.3846	4.6

Table 7. Results of condition 3 (audio and audio-haptic condition) for the gravel surface

	Bump	Hole	Flat	Mean certainty Total	Mean certainty Correct answers	Mean certainty Wrong answers	% Correct answers	Mean Realism	Mean Quality
flat_A			15	5.3333	5.3333	-	100	3.8	4.2
flat_AH	1	1	13	5.6667	5.8462	4.5	86.66	5.1538	5.2667
bump_step_A	12	2	1	4.5333	4.6667	4	80	3.75	4
bump_step_AH	12	2	1	5.3333	5.5833	4.3333	80	4.3333	4.8667
hole_step_A	1	14		5.1333	5.2857	3	93.33	4.2857	4.7333
hole_step_AH	1	13	1	5.0667	5	5.5	86.66	4.4615	4.9333
bump_h_t_A	12	1	2	4.1333	4.25	3.6667	80	2.9167	3.6667
bump_h_t_AH	12	1	2	5.4	5.25	6	80	4.5833	4.8667
hole_h_t_A		10	5	4.2667	4.4	4	66.66	2.8	3.7333
hole_h_t_AH	3	12		4.6667	4.75	4.3333	80	4.3333	5.3333
bump_comb_A	14	1		5.1333	5.2143	4	93.33	3.9286	4
bump_comb_AH	12	3		5.1333	5.3333	4.3333	80	4.0833	4.6
hole_comb_A	1	13	1	4.9333	5.1538	3.5	86.66	4	4.5333
hole_comb_AH	2	13		5.7333	6.2308	2.5	86.66	4.7692	5.2667

As regards the degree of certainty of the correct answers, most of the times conditions with audio-haptic modality present higher values than the audio modality, Despite an in-depth analysis with t-test did not show any statistically significant difference between the corresponding pairs, the ANOVA analysis revealed that at global level the AH condition gives rise to higher evaluation in the case of gravel material (p-value = 0.0449).

As concerns the degree of realism and the degree of quality, results are always higher for the bimodal stimuli rather than the unimodal ones (see figure 4). The ANOVA analysis shows that such differences are statistically significative, (p-value < 0.0001 and p-value < 0.0001 for realism in the conditions wood and gravel respectively, and p-value = 0.002862 and p-value = 0.003386 for quality in the conditions wood and gravel respectively). In addition an in depth analysis performed with t-test, revealed also significant differences for the pairs of stimuli (audio, audio-haptic) in table 8.

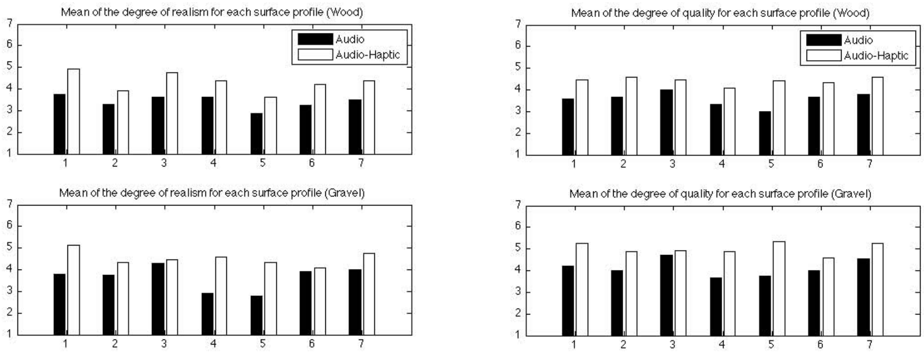

Fig. 4. Average scores of the degree of realism (left) and quality (right), for audio and audio-haptic conditions in condition 3. Surface profile from left to right: 1-flat, 2-bump_step, 3-hole_step, 4-bump_h_t, 5-hole_h_t, 6-bump_comb, 7-hole_comb.

Table 8. Statistically significative differences for average scores of the degree of realism and quality, between audio and audio-haptic conditions in condition 3

Stimulus	**Material**	**P-value realism**	**P-value quality**
flat	wood	0.03381	-
flat	gravel	0.01425	-
hole_h_t	wood	-	0.01834
bump_h_t	gravel	0.001482	0.04558
hole_h_t	gravel	0.02296	0.00732

Considering the total degree of certainty (column 5), it is possible to notice that on average the stimuli involving only the variations in the heel-toe distance present lower values in comparison with the other conditions. This is an indication of a greater difficulty in the participants' choices for these stimuli rather than the others.

Finally, an ANOVA analysis performed on the evaluations of the degree of certainty of correct answers, of realism and of quality, did not reveal any significant difference between the two materials.

5 General Discussion

From a comparison between results of the three condtions it is possible to notice that the use of the proposed techniques allows to simulate the act of walking on a bump, a hole, or a flat surface by passively presenting audio, haptic and audio-haptic stimuli. The most important result is that the haptic modality plays an important role in this kind of tasks. In particular, results of condition 3 show that haptic cues significantly reinforces the auditory cues. Indeed, participants gave rise to higher evaluations in presence of audio-haptic stimuli rather than the audio-stimuli, for what concerns the realism and the quality of the presented stimuli, as well as they were more certain in giving their correct answers. When the haptic is presented alone it is still possible to simulate quite well the act of walking over a bump, a hole or a flat surface, as results of condition 2 show. From a comparison with the results of condition 1 what emerges is that the auditory modality seems to be dominant on the haptic one in this kind of tasks. A common trend noticed in all the conditions is that the material does not have any particular influence on the participants evaluations, i.e., there is no statistically significant difference in the responses between the presented solid and aggregate surfaces. Another trend more or less common to the three conditions is that on average the stimuli involving only the variations in the heel-toe distance present lower values in comparison with the other conditions for what concerns the total degree of certainty, and this is an indication of a greater difficulty in the participants' choices for these stimuli rather than the others. This also means that participants found more natural the variations in the temporal distance between steps rather than the ones in the temporal distances between heel and toe. In addition, in all the conditions the combination of the two temporal information did not significantly enhance the recognition of a bump or a hole. Subjects were allowed to listen to the simulations as many times as they wished, but in general they listened only once, and this was enough to provide an answer. At the end of the test some informal interviews were performed, where subjects declared that the stimuli were clear. Subjects were also surprised by the ability of the feet to distinguish the simulations.

6 Conclusion and Future Work

In this paper, we described a between subject audio-haptic experiment whose goal was to assess the role of temporal aspects in recognizing whether a person is walking on a flat surface, a bump or a hole.

Results show that varying temporal aspects between footsteps allow to successfully simulating the act of walking on a bump, hole, or flat surface, especially in the auditory modality. Moreover, in the recognition task, haptic cues show to significantly reinforce auditory cues, as the results of condition 3 show. All three conditions were run by using the synthesis engine offline. The reason why we asked subjects to seat was to create an environment similar to a home situation, e.g., when subjects are watching a movie. Indeed, we plan to use the same engine with visual feedback, in order to assess the role of haptic feedback in enhancing the experience of watching an audio-visual scene.

In the future, we are interested in running the same experiments interactively, to understand whether our technique can be used to simulate bumps and holes in a virtual reality setup.

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.[4]. The authors

[4] Natural Interactive Walking Project: `www.niwproject.eu`

would like to thank Professor Vincent Hayward and Amir Berrezag for providing the haptic shoes used in this experiment.

References

1. Cook, P.: Physically Informed Sonic Modeling (PhISM): Synthesis of Percussive Sounds. Computer Music Journal 21(3), 38–49 (1997)
2. Cook, P.: Modeling Bill's Gait: Analysis and Parametric Synthesis of Walking Sounds. In: Proceedings of the AES 22nd International Conference on Virtual, Synthetic, and Entertainment Audio, pp. 73–78 (2002)
3. Farnell, A.: Marching onwards: procedural synthetic footsteps for video games and animation. In: Proceedings of the Pure Data Convention (2007)
4. Fontana, F., Bresin, R.: Physics-based sound synthesis and control: crushing, walking and running by crumpling sounds. In: Proc. Colloquium on Musical Informatics, pp. 109–114 (2003)
5. Hollerbach, J., Checcacci, D., Noma, H., Yanagida, Y., Tetsutani, N.: Simulating side slopes on locomotion interfaces using torso forces. In: Haptic Symposium, Citeseer, pp. 91–98 (2003)
6. Hollerbach, J., Mills, R., Tristano, D., Christensen, R., Thompson, W., Xu, Y.: Torso force feedback realistically simulates slope on treadmill-style locomotion interfaces. The International Journal of Robotics Research 20(12), 939 (2001)
7. Hunt, K.H., Crossley, F.R.E.: Coefficient of restitution interpreted as damping in vibroimpact. ASME Journal of Applied Mechanics 42(2), 440–445 (1975)
8. Iwata, H., Yano, H., Nakaizumi, F.: Gait master: A versatile locomotion interface for uneven virtual terrain. vr, page 131(2001)
9. Lécuyer, A., Burkhardt, J., Etienne, L.: Feeling bumps and holes without a haptic interface: the perception of pseudo-haptic textures. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, page 246. ACM, New York (2004)
10. Marchal, M., Lecuyer, A., Cirio, G., Bonnet, L., Emily, M.: Walking Up and Down in Immersive Virtual Worlds: Novel Interactive Techniques Based on Visual Feedback. In: Proceedings of IEEE Symposium on 3D User Interface (2010)
11. Miner, N., Caudell, T.: Using wavelets to synthesize stochastic-based sounds for immersive virtual environments. ACM Transactions on Applied Perception 2(4), 521–528 (2005)
12. Nordahl, R., Serafin, S., Turchet, L.: Sound synthesis and evaluation of interactive footsteps for virtual reality applications. In: Proc. IEEE VR (2010)
13. 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. Eurohaptics Symposium on Haptics and Audio-visual Environments (2010)
14. Serafin, S., Turchet, L., Nordahl, R.: Extraction of ground reaction forces for real-time synthesis of walking sounds. In: Proc. Audiomostly (2009)
15. Serafin, S., Turchet, L., Nordahl, R.: Do You Hear A Bump Or A Hole? An Experiment on Temporal Aspects in the Recognition of Footsteps Sounds (2010)
16. Turchet, L., Marchal, M., Lécuyer, A., Nordahl, R., Serafin, S.: Influence of auditory and visual feedback for perceiving walking over bumps and holes in desktop VR. In: Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology, pp. 139–142. ACM, New York (2010)

17. 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)
18. Turchet, L., Nordahl, R., Serafin, S.: Examining the role of context in the recognition of walking sound. In: Proc. of Sound and Music Computing Conference (2010)
19. Turchet, L., Serafin, S., Dimitrov, S., Nordahl, R.: Conflictual audio-haptic feedback in physically based simulation of walking sounds. In: Proc. of Haptic Audio Interaction Design Conference (2010)
20. 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)
21. Welch, R., Warren, D.: Immediate perceptual response to intersensory discrepancy. Psychological Bulletin 88(3), 638 (1980)