Applied Acoustics 75 (2014) 59–66

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Semantic congruence in audio–haptic simulation of footsteps

Luca Turchet *, Stefania Serafin

Department of Architecture, Design and Media Technology, Aalborg University Copenhagen, A.C. Meyers Vænge 15, 2450 Copenhagen, Denmark

ARTICLE INFO

Article history:
Received 5 December 2012
Received in revised form 17 June 2013
Accepted 24 June 2013
Available online 3 August 2013

Keywords:
Audio
Haptic
Semantic congruence

ABSTRACT

In this paper we present an experiment whose goal is to investigate subjects' ability to match pairs of synthetic auditory and haptic stimuli which simulate the sensation of walking on different surfaces. In three non-interactive conditions the audio–haptic stimuli were passively presented through a desktop system, while in three interactive conditions participants produced the audio–haptic feedback interactively while walking. Results show that material typology (i.e., solid or aggregate) is processed very consistently in both the auditory and haptic modalities. Subjects expressed a higher level of semantic congruence for those audio–haptic pairs of materials which belonged to the same typology. Furthermore, better matching ability was found for the passive case compared to the interactive one, although this may be due to the limits of the technology used for the interactive haptic simulations.

1. Introduction

In the field of multimodal perception and cognition, several studies have investigated the reactions of subjects when presented simultaneously with stimuli in two different modalities. Studies on the audiovisual multisensory integration show that the binding of the two modalities depends on several factors [1]. One of those is whether the stimuli are semantically congruent or not [2,3]. This means that subjects are able to provide a congruent meaning to the stimuli in both modalities. An example of a semantically congruent stimulus in the audio-visual domain is seeing the picture of a dog and hearing a barking sound. The same bimodal stimulus would be semantically incongruent if the picture of a dog was shown while hearing the sound of a cat meowing [2].

Semantically congruent stimuli have been shown to enhance behavioral performance [3,2], and to aid the identification of masked images [4]. However, further investigations are needed to determine under which conditions semantic congruency influences audiovisual multisensory integration [5].

On the other hand, several studies have been conducted on multimodal perception involving the auditory and haptic modalities [6–9]. However, to our knowledge few previous studies investigated the semantic congruence between audition and touch. This is especially the case when auditory and tactile stimuli are presented at feet level since research on the interaction between touch and audition has focused mainly on the hand [8,9].

Although the foot–ground interactions are phenomena which produce rich sensory information, few studies on both the auditory and haptic perception have been conducted in this context. At auditory level, Li and colleagues investigated the ability of subjects to identify the gender of a human walker by listening his/her footstep sounds [10], while Pastore and co-workers investigated listeners' ability to make judgments about the posture of the walker who generated the acoustic stimuli [11].

Moreover, the haptic perceptual system has been proven to be able to discriminate grounds of different elasticity while walking [12], and the vibrotactile sensory channels showed to play an important role in the perception of ground surface compliance during walking [13].

Furthermore, the interaction of auditory and haptic feedback in foot has been studied in [14], showing that the feet were also effective at probing the world with discriminative touch, with and without access to auditory information.

Recently, we developed an interactive system which can provide combined auditory and haptic sensations that arise while walking on solid and aggregate surfaces (the latter being assumed to possess a granular structure, such as that of gravel). The system is composed of an audio–haptic synthesis engine, and a pair of shoes enhanced with sensors and actuators able to provide plantar cutaneous vibration feedback. Such system can be used also non-interactively, providing the user with the audio–haptic feedback while sitting on a chair.

The ecological validity of the auditory as well as of the haptic stimuli involved in the present experiment was assessed in previous research. The results of two non-interactive listening experiments showed that the majority of the simulated surfaces was recognized with high accuracy [15,16]. In particular they were proven to be correctly classified in the corresponding solid and aggregate surface typology. Similar accuracy was noticed in the

* Corresponding author.
E-mail addresses: tur@create.aau.dk (L. Turchet), sts@create.aau.dk (S. Serafin).

0003-682X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.apacoust.2013.06.016

recognition of real recorded footstep sounds, which was an indication of the success of the proposed algorithms and their control. Analogously, results of a haptic recognition experiment involving the proposed non-interactive system indicated that subjects were able to correctly discriminate the typology of the simulated surfaces [16]. Moreover, the proposed non-interactive haptic feedback has been recently proven to significantly increase the realism of the desktop-simulated walking experience [17].

Furthermore, the ecological validity was also assessed for the interactive simulations, yielding results similar to those reported for the non-interactive case [15,18,17].

Moreover, from a comparison between the results of the auditory and haptic conditions in [16] and [18], it is possible to notice that participants were able to correctly categorize in both modalities the typology of the simulated surface materials (i.e., solid or aggregate). Indeed solid materials were rarely confused with aggregates, and vice versa, and participants, when not recognizing the presented surface material tended to classify it as another belonging to a same typology (e.g., wood–concrete, snow–frozen snow, dry leaves–forest underbrush) rather than to different typologies (e.g., wood–gravel, metal–dry leaves).

Similar results were found in a recognition task involving the walking on real materials, proving the ability of humans to distinguish almost perfectly between solid and aggregate materials both at auditory and haptic level [14].

Consequently, all these findings suggest that material typology is processed very consistently in the two modalities. In order to investigate the extent to which the two modalities are similar, we conducted an experiment in which we assessed the capacity of subjects to associate the auditory and foot-haptic stimuli provided both in semantically congruent and incongruent way, and both interactively and non-interactively.

To the best of our knowledge, the study of the semantic congruence between the auditory and foot-haptic modalities is a research topic still unexplored. The present research is relevant for the study of the multisensory perception of material properties during walking, for the understanding of the mechanisms underlying the multisensory categorization, and for the comparison of the structure of the perceptual spaces of the audition and foot-touch modalities.

2. Method

We conducted a between-subjects experiment divided in six conditions whose goal was to investigate the ability of subjects to match the different sounds and haptic sensations they were presented with. On three of the six conditions, subjects were not walking but were sitting on a chair and received passively the audio–haptic stimuli (non-interactive conditions). Conversely, in the other three conditions subjects produced the audio–haptic feedback interactively while walking (interactive conditions). The between subjects approach was chosen in order to avoid possible learning effects.

2.1. Apparatus

In previous research we developed a system which simulates both non-interactively and interactively the auditory and haptic sensation of walking on different surfaces [19,20]. To this purpose, shoes enhanced with actuators and pressure sensors were developed. The shoes were a pair of light-weight sandals (Model Arpenaz-50, Decathlon, Villeneuve d'Ascq, France). This particular model was chosen since it has light, stiff foam soles where it is relatively easy to insert sensors and actuators. Four cavities were made in 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 [21]. In each shoe, two actuators were placed under the heel and the other two under the toe. They were bonded in place to ensure good transmission of the vibrations inside the soles. When activated, vibrations propagated far in the foam. In addition, the sole had two force sensitive resistors sensors intended to pick the foot-floor interaction force in order to drive the auditory and haptic synthesis. The two sensors were placed in correspondence to the heel and toe respectively in each shoe.

The involved hardware allowed the control in real-time of an audio–haptic synthesis engine based on physical models [19]. For the purpose of the experiment, the engine was set in order to synthesize footstep sounds on both solid and aggregate materials, which were simulated using an impact model [22] and a physically informed sonic model (PhiSM) algorithm [23]. In particular, the same models were used to drive both the haptic and the auditory synthesis. They are briefly recalled below.

In the simulation of impact with solids, the contact was modeled by a Hunt–Crossley-type interaction where the force, f, between two bodies, combines hardening elasticity and a dissipation term [24]. Let x represent contact interpenetration and $\alpha > 1$ be a coefficient used to shape the nonlinear hardening, the special model form we used is:

$$f(x,\dot{x}) = -kx^{\alpha} - \lambda x^{\alpha}\dot{x} \quad \text{if } x > 0, \quad 0 \text{ otherwise}. \tag{1}$$

The model described was discretized as proposed in [22].

To simulate aggregate surfaces, the PhiSM algorithm was adopted. This algorithm simulates particle interactions by using a stochastic parameterization thereby avoiding to model each of many particles explicitly. Instead, the particles are assigned a probability to create an acoustic waveform. In the case of many particles, the interaction can be represented using a simple Poisson distribution, where the sound probability is constant at each time step, giving rise to an exponential probability weighting time between events.

In the interaction between a foot and a sole an important element is the ground reaction force (GRF), i.e., the reaction force supplied by the ground at every step. In our simulations the physical models were driven by a signal expressing the GRF, which corresponded to the amplitude envelope extracted from an audio signal containing a footstep sound.

The synthesis engine can work both interactively and non-interactively. In this study to control the engine interactively, we used recorded GRF files corresponding to heel and toe strikes which were triggered according to the shoe sensor activated during the user locomotion [20]. To control the engine non-interactively, we created different audio files using the recording of a single real footstep sound on concrete. The envelope profiles of each step in the file were extracted and fed to the engine which produced the synthesized footstep sounds according to the choice of the surface to be simulated. In particular, the surface profile chosen for the experiments was the flat one. It was simulated by placing the footstep sound generator at equal intervals of time, precisely 750 ms, which corresponds to a moderately fast gait [25].

2.1.1. Setup

The experiment was carried out in an acoustically isolated room where the setup was installed. It consisted of a simple graphical user interface with which participants were asked to interact, a spreadsheet to collect their answers, a set of headphones (Sennheiser HD 600) and the haptic shoes previously described. The interface comprised numbered buttons. A button press triggered the presentation of the audio–haptic stimulus through the headphones