Rhythmic walking interaction with auditory feedback

Ecological approaches in a tempo following experiment

Maculewicz, Justyna; Jylhä, Antti; Serafin, Stefania; Erkut, Cumhur

Published in:
IEEE MultiMedia

DOI (link to publication from Publisher):
10.1109/MMUL.2015.27

Publication date:
2015

Document Version
Accepted author manuscript, peer reviewed version

This article has been accepted for publication in IEEE MultiMedia but has not yet been fully edited.
Some content may change prior to final publication.

Rhythmic Walking Interaction with Auditory Feedback: Ecological Approaches in a Tempo Following Experiment

Justyna Maculewicz[1], Antti Jylhä[2], Stefania Serafin[1], and Cumhur Erkut[1]

[1]Department of Architecture, Design and Media Technology, Aalborg University Copenhagen, Denmark

[2]Helsinki Institute for Information Technology (HIIT), Department of Computer Science, University of Helsinki, Finland

Abstract

We present an interactive auditory display for walking with sinusoidal tones or ecological, physically-based synthetic walking sounds. The feedback is either step-based or rhythmic, with constant or adaptive tempo. In a tempo-following experiment, we investigate different interaction modes and auditory feedback, based on the MSE between the target and performed tempo, and the stability of the latter. The results indicate that the MSE with ecological sounds is comparable to that with the sinusoidal tones, yet ecological sounds are considered more natural. Adaptive conditions result in stable tempi. These results have implications on the design of interactive entertainment or therapeutical applications.

Keywords: Audio input/output; rhythmic interaction, walking, ecological feedback

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

This article has been accepted for publication in IEEE MultiMedia but has not yet been fully edited.
Some content may change prior to final publication.

1 Introduction

Walking is an activity that plays a very important part in our daily life. In addition to being a natural means of transportation, walking is also characterized by the resulting sound, which can provide information about the surface, type of shoes, and movement speed as well as the person's age, weight, and physical condition.

Lately, there have been several research and development efforts in different fields related to walking. In the commercial world, interactive system developers have become increasingly interested in new interaction paradigms based on the feet. As a notable example related to walking, a set of console games have been specifically designed making use of controllers such as the Nintendo Wii Fit balance board as a tracking device for the gait of the player. This tracking is directly based on sensing the force imposed by the human on the board while standing or walking *in place*. The feedback in these games is varied but typically contains auditory and visual elements that react to the actions of the player. For example, the player may need to walk fast on the board to pedal a virtual bicycle or to march to the rhythm of a reference music, making the avatar walk steadily.

In the field of interaction design, interest in the use of feet as a mode of interaction has recently increased. As an example, the Natural Interactive Walking (NIW) EU project, which ended in 2011, investigated possibilities for the use of the auditory and haptic modalities in floor-based interfaces, and for the synergy of perception and action in capturing and guiding human walking in *locomotion*. A goal of the NIW project was to provide closed-loop interaction paradigms validated through experiments, enabling the transfer of skills that have been previously learned in everyday tasks associated to walking. During the project, walkers were exposed to virtual scenes presenting grounds of different natures in which to situate the sensing and display of haptic and acoustic information for interactive simulation. Different experiments measured the ecological validity of such scenarios, investigating also the cognitive aspects of the underlying perceptual processes [17, 13, 14]. Note

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

that, by 'ecological validity' we refer both to the real-world applicability as used in the experiment design, and to the correlation between the sensory cue (in our case, walking sounds) and the walker as defined in ecological psychology.

Furthermore, in medical sciences, walking exercises have emerged as means for rehabilitation on treadmill or floor. Indeed, studies have investigated the possibility to benefit gait patterns by using rhythmical auditory cues, especially for Parkinson patients [8, 16]. Typically, in these studies cue signals have been obtained from digital metronomes, which have low ecological validity.

Most of the walking interfaces and interaction paradigms of the recent past have been focusing on a one-to-one mapping between the steps of the user and the feedback, and have not examined any rhythmical patterns and continuous temporal evolutions of the action of the player. In addition, while the differences between the metronome and music pacing have been identified [15, 11], the ecological sounds are still a grey area yet to be researched. Finally, tempo and rhythms are less frequently used in sonification compared to other auditory attributes [3]. The effect of ecological auditory feedback on rhythmic walking therefore poses interesting questions for interactive sonification.

In this paper, we study walking as a rhythmic action and experimentally investigate the effect of auditory feedback on this action, within the framework of closed-loop interactive sonification [7]. The study builds on a pilot experiment [10] with a broader study and new experimental conditions. The subjects were provided with different auditory cues to guide their gait at different tempi. As the input modality tracking the footsteps, we use audio, captured near the feet with a condenser microphone.

Our aim is to study how different rhythmic cues affect the walking rhythm, thus providing insight into the design of rhythmic feet-based interactive systems. We consider different temporal forms of the feedback, namely direct synthetic response to each step, and both natural and unnatural

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

This article has been accepted for publication in IEEE MultiMedia but has not yet been fully edited.
Some content may change prior to final publication.

continuous synthetic audio feedback with and without tempo adaptation to the human gait. We study the differences in performance and user experience between these experimental conditions.

In the next section the background of the study is discussed. Section 3 presents the experiment setup and procedure, after which results are presented and discussed in Sections 4 and 5, respectively. Section 6 concludes the paper.

2 Background

The study of walking as a rhythmic activity has been quite widely investigated in the research community. In significant research of Stynes et al. [15] participants were asked to synchronize their walking tempo with the tempo of musical and metronome stimuli. Their results show a relationship between rhythm perception and locomotion: people seem to walk faster to music than to metronome stimuli. More recently, a large collection of motivational music was equalized to 130 beats per minute (BPM), and the vigor of walking was found to be related to beat salience through movement imagery [11]. In other words, if people clearly perceive the beat structure in music, this perception entrains how they imagine walking to it, and results in observable change in walking qualities. A similar relation in ecological walking sounds has not been investigated, to our knowledge. However, our everyday experience indicates that, for example, walking on gravel requires different vigor compared to walking on wood, which may be entrained by specific walking sounds. To test this relationship, we need a fast and reliable means of capturing and resynthesizing walking sounds, and to combine them in rhythmic interaction.

2.1 Sound as an Input Modality

In our study, we use walking sounds as input to the system. One of the advantages of using sound instead of other sensor modalities, like sensor- or video-based tracking, is the satisfactory temporal

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

This article has been accepted for publication in IEEE MultiMedia but has not yet been fully edited.
Some content may change prior to final publication.

resolution with non-invasive sensors and lower latency. The typical sampling frequency for audio is 44100 Hz, whereas for most sensor-based devices such as accelerometers it is at most 100 Hz. Using sound as input to interactive applications, however, requires efficient algorithms for processing and understanding the audio in order to avoid latency.

Real-time techniques for retrieving various kinds of information from percussive sounds have been proposed in different contexts. As an example, in the context of rhythmic hand clapping, the recognition of the hand clapping type and tempo have been utilized in construction of interactive systems using sound as the principal input modality [9]. Rhythmic interaction between a human clapper and the computer has been studied by a virtual flamenco rhythm tutor application that has also been used as a tool for evaluating rhythmic interactive systems in general. The findings show, for example, that in rhythmic learning auditory feedback is crucial and that slight tempo adaptation of the tutor improves the performance [9].

In this paper we are interested in investigating how people react when exposed to synthetic footstep sounds. Synthesis of footstep sounds has been investigated from several perspectives, described in the following section.

2.2 Synthesis of Walking Sounds

Several algorithms have been proposed to simulate the sound of walking. One of the pioneers in this domain is Perry Cook, who proposed a collection of physically informed stochastic models (PhISM) simulating several everyday sonic events [1]. Among such algorithms were simulated the sounds of people walking on different surfaces [2]. A similar algorithm is also proposed in [6], where physically informed models simulate several stochastic surfaces. The characteristic events of footstep sounds were reproduced by simulating the so-called ground reaction force, i.e., the reaction force supplied by the ground at every step.

This article has been accepted for publication in IEEE MultiMedia but has not yet been fully edited.
Some content may change prior to final publication.

2.3 Rhythmic Interaction

Rhythmic interaction refers to an interactive experience in which the rhythms are explicitly felt as an interaction attribute [4]. A challenge in this interaction is the alignment of the input and output modalities within ecologically-correct settings. Especially when an output modality that resembles a self-produced sound, such as synthetic hand claps or walking sounds is aligned with the rhythmic input, we have argued that this could provide a tighter interaction loop and strengthen the feeling of *agency* [4]. Agency, i.e., the feeling of being aware and in control of individual actions in the presence of multimodal feedback, has been traditionally investigated and evaluated by using upper limb movements. When tested in the context of walking, agency judgments show a complex modulation depending on the natural stride-time of the subjects and the feedback delay of the system (see the references in [4]). Subjects also change their stride-time and alter their gait parameters. These observations indicate that rhythmic interaction with realistic auditory feedback has a great potential in various applications of interactive sonification, such as in gait rehabilitation.

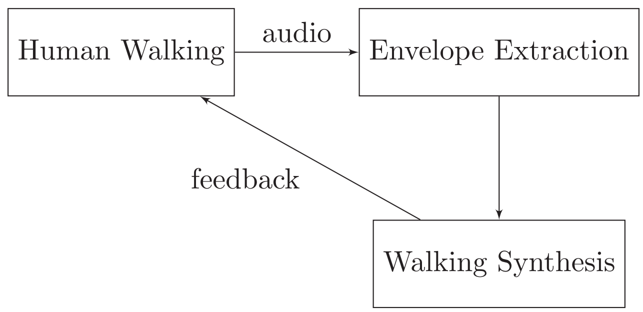
3 Experiments

3.1 System Architecture

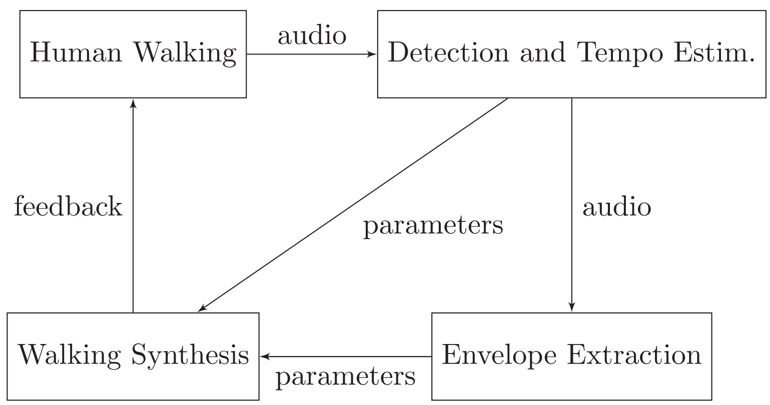
To examine rhythmic walking interactions we have constructed a system consisting of a synthesis engine for walking sounds and two different detection modules for tracking the footsteps of the user [10]. The principles of the synthesis engine are described in more detail in [13, 10] . The system reproduces synthetic walking sounds based on extracting the envelope of human footsteps and using it to drive a physically informed synthesis model. This setting directly synthesizes a footstep sound preserving the envelope of the step of the user. In an alternative setting, the footsteps of the human are detected as events and the walking tempo is estimated. This information is used to control the

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

synthesis engine by sequentially triggering a footstep sound according to the estimated tempo. In this setting, the system thus produces footsteps as stimulus, either with a fixed tempo or with one adapting to the tempo of the human walker. This way, we get three temporally different conditions for auditory feedback. The data flow for each setting is presented in Figure 1.



a)



b)

Figure 1: Information flow illustration of the system in two modes of function: a) direct synthesis driven by the envelope of human walking; b) the steps and tempo of the human walker are detected and used to inform the synthesis by sending tempo event parameters besides the envelope parameters.

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

This article has been accepted for publication in IEEE MultiMedia but has not yet been fully edited.
Some content may change prior to final publication.

E#	I-MODE	FEEDBACK	TEMPO (BPM)	TASK
a1	STEP (Silent)	None	-	Walk at own pace
a2	STEP (Sonified)	Sine, wood, gravel	-	#
b1	TEMPO (Constant)	Sine, wood, gravel	80, 105, 130	Follow tempo presented by sound
b2	TEMPO (Adaptive)	Sine, wood, gravel	80, 105, 130	#

Table 1: Summary of our experiment design. For experiment and interaction modes on the first and second columns, respectively, refer to Fig. 1.

3.2 Interaction Modes

To investigate the effect of different feedback sounds and interaction types on human walking, we designed an experiment where we compared three different interaction modes, three kinds of auditory feedback, and three tempi. The three interaction modes were 1) instantaneous auditory feedback per detected step, 2) pacing with a constant tempo, and 3) pacing with a tempo that adapts to the instantaneous walking pace of the subjects. In all interaction modes, auditory feedback was one of the following: I) a 1 kHz sinusoidal, II) synthetic footstep sounds on wood, and III) on gravel. The tempo was chosen among 80, 105, and 130 BPM, each beat corresponding to one step. These were chosen to be below, around, and above the typical walking pace of humans. Table 1 summarizes our experiment design.

3.3 Technical Apparatus and Setup

Figure 2 shows the developed system together with a user. The footstep sounds produced by the person walking in place on a step machine are captured by a microphone (Shure BETA 91[1]). We adopted a solution based on a step machine since the goal of this experiment was to emulate a typical home scenario, where devices such as the Wii Fit by Nintendo are present. This creates a

[1] http://www.shure.com/

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

This article has been accepted for publication in IEEE MultiMedia but has not yet been fully edited.
Some content may change prior to final publication.

situation which resembles the well-known walking in place techniques used in virtual reality research [12].

Figure 2: Left: The feet of a user interacting with the previous setup in [10]. Right: A single microphone is placed under the surface of the platform used in the new experiment.

The real-time analysis algorithms illustrated on Fig. 1 extract some salient features from such footsteps [10]. Specifically, the microphone is placed under an aerobic step. We chose a high performance condenser microphone with a tailored frequency response designed specifically for kick drums and other bass instruments. Its features make it a good candidate for capturing footstep sounds. The single microphone gave comparable results to the two-microphone system we have used before in [10], and we preferred this setup as it reduces the data size by a factor of two. The synthesized sounds are conveyed to the user by means of an open-design headphone. Note that contact microphones could be good alternatives to replicate our experiments in other walking

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

conditions, where a quite environment is not warranted or a multichannel loudspeaker system is preferred over a headphone as an auditory display for conveying sounds.

3.4 Experiment Sections

An experiment was designed to examine three different feedback conditions (interaction modes) and the way subjects respond to them while varying the feedback sounds and the tempi, with the aim of finding how the conditions affected the rhythmic behavior. It consisted of three parts and two questionnaires.

3.4.1 Step-wise Interaction

In the first part, the task of the participants was to walk for around one minute in their own preferred pace (test case a). This part had two sessions. The first session a1 was without auditory feedback. In the second session a2, the detected envelope of the footstep sounds of the subjects was applied as the driver for synthetic footstep sounds. The control variables in this case were the walking sounds on gravel and wood, and a sinusoidal sound (1000 Hz). The sinusoidal was obtained by tuning one of the model resonators to the desired frequency and amplitude. After three takes with auditory feedback, which were presented in random order, the participants were asked how much they agree with the statement “The sound of footsteps sounded as a natural consequence of walking”, on a scale of five, where the extremes pointed strong disagreement (1) or agreement (5).

3.4.2 Tempo-wise Interaction

The second part of the experiment (test case b) consisted of nine 40 seconds-long sessions each for b1 and b2. In both sessions, the participants had the same task, to follow the tempo presented by the sound. The difference between two sessions was that the level of the presented tempo in b1 was constant and in b2, it was adaptively adjusting to the detected tempo of the participant,

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

which could influence the decrease or increase in participant's tempo. The adaptation was limited to $\pm$ 1 % of the instantaneous tempo, which was considered undetectable by humans in the range we are interested in. For all the test cases, the sounds of walking were recorded, along with a log of time stamps for each detected step (see Section 3) as well as the continuously updated tempo estimate. Online tempo estimate was computed from the moving average of six previous onset-to-onset interval values.

3.4.3 Types of Feedback and the Tempi Levels

The number of sessions was a result of combination of three different kinds of guiding sounds and three different tempi. In a2, b1, and b2, the auditory feedback was gravel, wood or sine wave. The tempi were 80, 105 and 130 BPM in part b. We assumed that 105 BPM would be the most comfortable tempo for most of the participants to walk. Then, we chose 80 and 130 BPM to be significantly below and above the average pace. Our choices are comparable to those presented in previous experiments on walking with a metronome and music [11].

After each session in part b, the participants were asked to what extent they agreed with the statement: "It was easy to follow the tempo", on a scale of five similar to that described above.

3.4.4 Questionnaires

Two questionnaires after b1 and b2 consisted of questions about the naturalness of the presented sounds, the tempo preferences and the ease in following the tempo with specific feedback. Finally, the participants provided information about themselves, including age, gender, and self-reported musical skills, as customary in similar tests [11].

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

This article has been accepted for publication in IEEE MultiMedia but has not yet been fully edited.
Some content may change prior to final publication.

4 Results

Twenty subjects (8 females and 12 males) participated in the experiment. Their average age was 30.5 (s=11.5). In this section, the results of the experiment are presented, focusing on the control variables and the answers to the questionnaire.

The analysis of experimental data was based on information about the participant's natural walking tempo measured in BPM obtained from a1. The tempo was recorded every time the system detected one step. The results obtained at each session were stored in text files. The sixty last steps from each session were analysed. If a session shorter than 65 steps was detected, then all of the steps were analysed, excluding the first five.

The analysis was conducted based on the average tempi, the mean square error (MSE) between the given (target) and performed (measured) tempo, and the trend of the tempo changes obtained from the slope of a line fitted to the measured tempo. Fig. 3 depicts these variables in two cases. The sign of the trend indicates deceleration or acceleration, and its value quantifies this in the experiment.

In analysis we have used four statistical tests: one-way repeated-measures ANOVA, factorial repeated-measures ANOVA for metric values and Friedman's ANOVA and Wilcoxon signed-rank test for ordinal data from questionnaires [5]. The use of the listed tests was dictated by the types of the data, which we obtained in the experiment, amount of factors and a type of the experimental plan. All test are for within group experimental plan. Repeated measures test is for data, which were obtained in experiment with more than one factor (independent variable).

4.1 Step-wise Interaction Results

Analysis of the data from part a showed that participants walked in the slowest tempo when they did not have any additional feedback in a1 ($mean = 101.1, s = 11.3$). Additional feedback

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

This article has been accepted for publication in IEEE MultiMedia but has not yet been fully edited.
Some content may change prior to final publication.

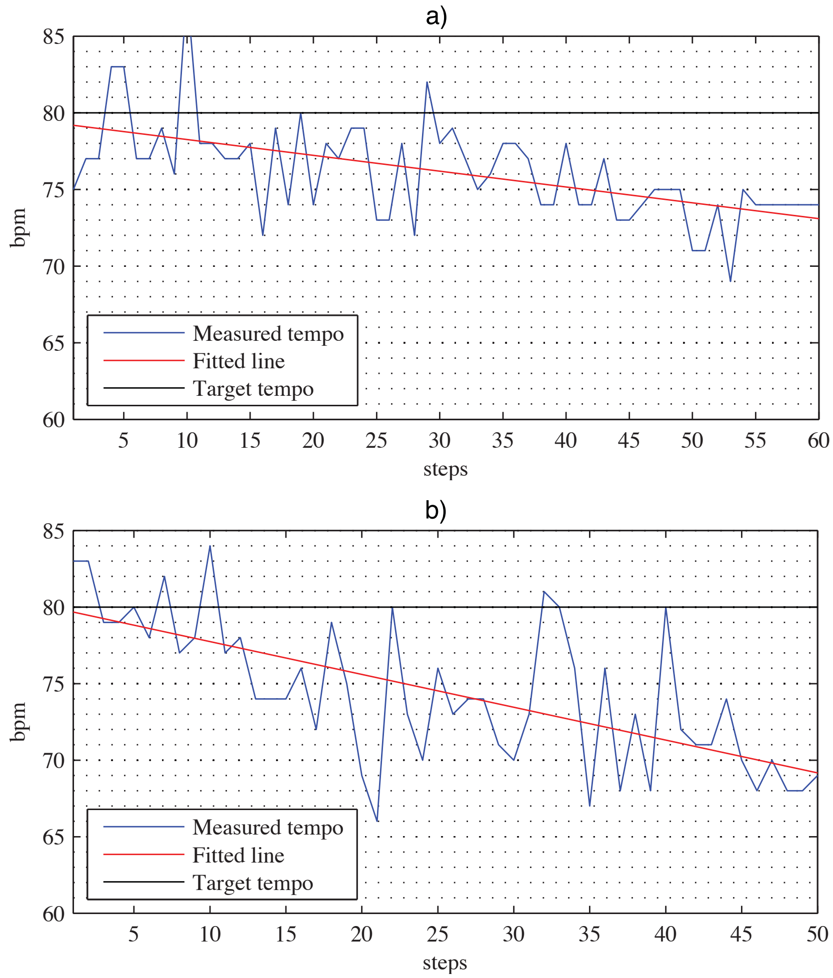


Figure 3: Two cases illustrating the analysis variables in 80 BPM constant tempo with wood sound feedback. a) Subject 7 and b) Subject 12.

in a2 changed their behaviour. Feedback, which imitated gravel sound made them walk faster ($mean = 103.2, s = 10.2$) than without feedback. The wood ($mean = 108.3, s = 11.6$) and the sine wave ($mean = 109, s = 11.5$) feedback induced even faster tempo of free walking. Statistical analysis (One-way repeated measures ANOVA) revealed that type of feedback had a significant effect on the average participants' tempo, $F(2.4, 2553.4) = 402.6, p < 0.001$.

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

This article has been accepted for publication in IEEE MultiMedia but has not yet been fully edited.
Some content may change prior to final publication.

4.2 Tempo-wise Interaction Results

Analysis of Mean Square Error (MSE) and trend in tempo changes showed that the type of feedback ($F(2, 1.9) = 7.3, p = 0.002$), the tempo ($F(1.2, 23.6) = 36.6, p < 0.001$) and the method (constant or adaptive) ($F(1, 19) = 21.6, p < 0.001$) significantly differentiated the MSE value in part b (Fig. 4). Factorial repeated measures ANOVA was used for these data. During the sessions with gravel feedback participants produced the biggest MSE, which indicates how much the participants deviated from the suggested tempo. The MSE was comparably smaller in the other two cases, slightly in favor of wood. Different tempo levels also caused different values of MSE. In the 80 BPM sessions the MSE value was the lowest and in 130 BPM the highest.

The trend in tempo changes indicates either acceleration (positive values) or deceleration (negative values). Analysis of the trend showed dependencies on tempo ($F(2, 38) = 8.4, p = 0.001$) and the feedback method ($F(1, 19) = 15.5, p = 0.001$) (Fig. 5). Acceleration has only been observed in constant tempo condition with the wood sounds, whereas the deceleration was also the highest for this feedback in adaptive conditions. Gravel trends in both cases were between that of the wood and the sine wave. In 80 BPM and 130 BPM sessions, the participants were slowing down and in 105 BPM they were accelerating. Finally, the absolute value of the trend was lowest in the 105 BPM sessions.

4.3 Questionnaire Results

During a2, the participants were asked if the sound of footsteps was perceived as a natural consequence of walking. Results indicated the gravel ($mean\ rank = 3.55$) as the most natural, followed by the wood ($mean\ rank = 2.65$). The sine wave ($mean\ rank = 1.9$) was the least natural. Friedman's ANOVA revealed that ratings significantly differed between the three kinds of feedback ($\chi^2 = 16.2, p = 0.001$).

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

This article has been accepted for publication in IEEE MultiMedia but has not yet been fully edited.
Some content may change prior to final publication.

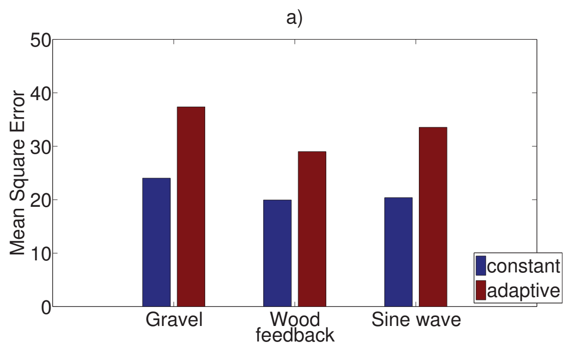


Figure 4: Relationship between the MSE and the type of feedback (a) and between the MSE and the tempo (b).

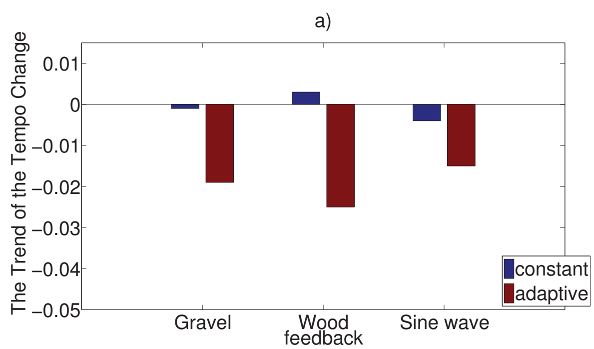


Figure 5: Relationship between the trend of the tempo change and the type of feedback (a) and between the tempo change and the tempo (b).

In the questionnaires after b1 and b2, the participants were asked if the gravel and wood feedback sounds were natural. Wilcoxon signed-rank test revealed that the statistical difference between these two types of feedback is significant, for the benefit of gravel sound ($z = -4.1, p < 0.001$).

After each session of part b, participants were asked to rate if it was easy to follow the tempo. Analysis showed significant difference between different tempi levels ($\chi^2 = 25.2, p < 0.001$). 105 BPM obtained the highest rating ($mean = 4.3$). The type of feedback did not differentiate ratings significantly. From this analysis, it is possible to conclude that wood was rated as the easiest type of feedback to follow, but ratings of different types of feedback were very similar.

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

5 Discussion

In general, feedback with ecological sounds resulted in a performance comparable to that with the sine wave, and in the case of wood, resulted in a more stable tempo. The realistic sounds were also considered more natural. In relation to hand claps, we have previously observed the benefits of contextual ecological sounds in tempo regularization and perceived naturalness of rhythmic interaction [4, 9]; the results reported here are a first step towards quantifying these in case of walking. The fact that these effects seem to be more pronounced for the wood sound is explained below, in relation to our everyday experience of walking on solid and aggregate surfaces.

Questionnaire data showed how the participants perceived the naturalness of feedback and also the ease of task execution. The gravel sound was found the most natural, followed by the wood and the sine wave. 105 BPM was the easiest to follow, explained by the fact that it was the most similar to the free walking tempo of the participants. The type of the feedback did not differentiate significantly the perceived ease of following the tempo, though we can say that wood obtained the highest rating.

This result is interesting, since sinusoidals are usually believed to provide fully predictable synchronization anchors, in contrast to our observations. It is likely in our case that the naturalness of the sounds might have interfered with the tempo following task. Similar to the beat salience in music [11], our experience of walking on solid and aggregate surfaces and familiarity with the resulting walking sounds may provide an advantage on steady tempo performance over sine tones. The GRF is much more spread in time and space on aggregate surfaces compared to solid surfaces, which may add to the perceived naturalness. The same spread, however, may modulate both the perceived beat salience from the rhythmic walking sounds and the preparatory motor action of the walker, and may explain why wood got the highest rating.

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

Results from the free walking and step sonification tasks showed that participants had the fastest tempo with wood. We believe this, too, is related to finding stable synchronization anchors: no feedback and gravel feedback conditions gave slower tempo, because participants had to focus more on finding anchors for synchronizing their performance to the auditory feedback when present, and keeping up with their internal clock when absent.

Analysis of the average tempi, the MSE, and the trends in acceleration/deceleration showed that in 105 BPM adaptive condition the participants were increasing their tempi. In contrast, both the 80 BPM and 130 BPM adaptive conditions resulted in deceleration. In the constant tempo condition, however, participants kept a steady tempo regardless of the type of auditory feedback. Their musical background did not have an effect on this observation. Interestingly, there was also no significant difference on their ratings related to the task difficulty.

Finally, our results show that the trend value (the slope of the measured tempo) was differentiated by the tempo and the absolute trend value was the lowest in 105 BPM conditions, which was found as the most natural tempo for participants in free walking. Around the preferred tempi, the participants were overall more stable. The 105 BPM condition resulted in a MSE value comparable with 80 BPM, but the absolute trend value was the lowest. The MSE was the highest for gravel conditions. It might be because it was the least specific kind of feedback in tempo punctuation. In addition, the choice of wood may also reflect the haptic perception of the step; an incongruent multimodal perception might be in effect when the subjects hear gravel but feel a solid, non-aggregate surface at their feet.

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

This article has been accepted for publication in IEEE MultiMedia but has not yet been fully edited.
Some content may change prior to final publication.

6 Conclusions

In this paper we describe a system that exploits the rhythmicity of footsteps in an interactive context. The system is composed of a microphone detecting footstep sounds, sound analysis and synthesis algorithms, and auditory feedback provided through headphones.

We have run an experiment in order to evaluate the way the auditory feedback affected people's walking characteristics. The specific focus was on the kind and different temporal forms of the feedback: direct synthesis of walking based on the amplitude envelope of the detected footstep, and continuous synthetic walking sounds with a fixed-tempo and an adaptive tempo. Multimedia examples demonstrating our experiments can be found at `http://www.smc.aau.dk/research`.

Concerning the feedback from participants, results show that subjects found the rhythmic interaction quite intuitive, and type of the feedback did not differentiate significantly the perceived ease of following the tempo. Their performance results, on the other hand, point out significant differences.

Results from the free walking and step sonification tasks showed that participants had the fastest tempo with wood. We tentatively conclude that this interaction falls in the grey area between the music and synthetic sounds of a metronome.

Between the conditions where the auditory feedback was provided with a constant and adaptive tempo, the participants did not report perceivable difference. In the constant tempo conditions, participants kept a rather steady tempo regardless of the type of auditory feedback. The adaptive condition resulted in a change: in 105 BPM adaptive condition participants were increasing their tempi, whereas both the 80 BPM and 130 BPM adaptive conditions resulted in deceleration.

Implications for design can be summarized as follows: both ecological sounds and an adaptive tempo have an effect on desired performance, and can provide additional design dimensions to experiment with in rhythmic interaction. Because of the effect of subtle temporal differences

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

in constant and adaptive conditions, the description of the exact interaction dynamics are of paramount importance. Finally, variability in auditory feedback may increase the motivational value of similar interactive systems.

In the future, we plan to use the system in a larger walking area and investigate signal detection algorithms that are also able to track the position where a person is walking. Moreover we plan to apply the system in interactive walking situations where rhythmic interaction plays an important role.

References

[1] P. Cook. Physically informed sonic modeling (PhISM): Synthesis of percussive sounds. *Computer Music J.*, 21(3):38–49, Summer 1997.

[2] P. Cook. Modeling Bill's gait: Analysis and parametric synthesis of walking sounds. In *Proc. AES 22nd Int. Conf. Virtual, Synthetic, and Entertainment Audio*, pages 73–78, June 2002.

[3] G. Dubus and R. Bresin. A Systematic Review of Mapping Strategies for the Sonification of Physical Quantities. *PloS one*, 8(12):e82491, 2013.

[4] C. Erkut, A. Jylhä, and D. Rocchesso. Heigh Ho: Rhythmicity in sonic interaction. In S. Serafin and K. Franinovic, editors, *Sonic Interaction Design*, chapter 17. The MIT Press, Boston, MA, USA, Spring 2013.

[5] A. Field. *Discovering statistics using IBM SPSS statistics.* Sage, London, UK, 2013.

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

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

[7] T. Hermann and A. Hunt. An introduction to interactive sonification. *IEEE Multimedia*, 12(2):20–24, 2005.

[8] T. Howe, B. Lövgreen, F. Cody, V. Ashton, and J. Oldham. Auditory cues can modify the gait of persons with early-stage Parkinson's disease: a method for enhancing parkinsonian walking performance? *Clinical Rehabilitation*, 17(4):363–367, July 2003.

[9] A. Jylhä, I. Ekman, C. Erkut, and K. Tahiroğlu. Design and evaluation of rhythmic interaction with an interactive tutoring system. *Computer Music J.*, 35(2):36–48, May 2011.

[10] A. Jylhä, S. Serafin, and C. Erkut. Rhythmic walking interactions with auditory feedback: An exploratory study. In *Proceedings of the Audio Mostly Conference*, pages 68–75, 2012.

[11] M. Leman, D. Moelants, M. Varewyck, F. Styns, L. Noorden, and J.-P. Martens. Activating and Relaxing Music Entrains the Speed of Beat Synchronized Walking. *PloS one*, 8(7):e67932, July 2013.

[12] N. C. Nilsson, S. Serafin, M. H. Laursen, K. S. Pedersen, E. Sikström, and R. Nordahl. Tapping-In-Place: Increasing the naturalness of immersive walking-in-place locomotion through novel gestural input. In *Proc. IEEE Symposium on 3D User Interfaces*, pages 31–38, Orlando, FL, USA, Mar. 2013. IEEE.

[13] R. Nordahl, L. Turchet, and S. Serafin. Sound synthesis and evaluation of interactive footsteps and environmental sounds rendering for virtual reality applications. *IEEE Trans. Visualization and Computer Graphics*, 17(9):1234–1244, September 2011.

[14] S. Papetti, F. Fontana, M. Civolani, A. Berrezag, and V. Hayward. Audio-tactile display of ground properties using interactive shoes. In *Proc. Int. Conf. Haptic and Audio Interaction Design*, pages 117–128. Springer, September 2010.

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE

[15] F. Styns, L. van Noorden, D. Moelants, and M. Leman. Walking on music. *Human Movement Science*, 26(5):769–785, October 2007.

[16] M. Thaut and M. Abiru. Rhythmic auditory stimulation in rehabilitation of movement disorders: a review of current research. *Music Perception*, 27(4):263–269, April 2010.

[17] Y. Visell, F. Fontana, B. Giordano, R. Nordahl, S. Serafin, and R. Bresin. Sound design and perception in walking interactions. *Int. J. Human-Computer Studies*, 67(11):947–959, November 2009.

Digital Object Indentifier 10.1109/MMUL.2015.27 1070-986X/$26.00 2015 IEEE