**Research Article**

Maximilian Schrapel*, Janko Happe, and Michael Rohs

# EnvironZen: Immersive Soundscapes via Augmented Footstep Sounds in Urban Areas

https://doi.org/10.1515/icom-2022-0020

**Abstract:** Urban environments are often characterized by loud and annoying sounds. Noise-cancelling headphones can suppress negative influences and superimpose the acoustic environment with audio-augmented realities (AAR). So far, AAR exhibited limited interactivity, e. g., being influenced by the location of the listener. In this paper we explore the superimposition of synchronized, augmented footstep sounds in urban AAR environments with noise-cancelling headphones. In an online survey, participants rated different soundscapes and sound augmentations. This served as a basis for selecting and designing soundscapes and augmentations for a subsequent in-situ field study in an urban environment with 16 participants. We found that the synchronous footstep feedback of our application EnvironZen contributes to creating a relaxing and immersive soundscape. Furthermore, we found that slightly delaying footstep feedback can be used to slow down walking and that particular footstep sounds can serve as intuitive navigation cues.

**Keywords:** soundscape, audio-augmented reality, augmented reality, auditory feedback, pedestrian, walking, mobile computing, urban computing

## 1 Introduction

Urban soundscapes are often far from what we consider relaxing. Car engines, horns, construction sites, laughing, or clanking dishes in restaurants and other sound sources can raise environmental noise to an annoying level. Research found a relation between environmental noise and various diseases such as tinnitus, cognitive impairment, sleep disturbance, and cardiovascular effects [48, 69, 74]. Animals also suffer from man-made noise [57], like birds singing at higher amplitude in urban areas [19]. The world health organization (WHO) therefore provides guidelines prescribing a maximum noise level [11, 75]. Human ears adapt to modern environments and focus on specific sound sources from any direction while attenuating other noises around.

Nevertheless, our natural ability to filter out distracting noises is limited [87]. Headphones with active noise cancellation (ANC) can effectively reduce distracting noises [68] and positively influence subjective well-being while focusing on specific tasks [70]. The listener can enjoy music or audio books at a lower volume [20], but may also miss important signals like announcements on trains [37]. Noise cancellation can also be beneficial for digitally supported meditation practice [91]. Walking meditation is a form of focusing on one's gait by paying attention to each step [39]. Research has found positive effects on the treatment of depression via walking meditation [77]. Typically, these exercises are performed in natural and relatively quiet environments to avoid distractions and to allow practitioners to better focus on the moment [52]. For novices, it is challenging to transform a walk on a crowded and noisy sidewalk into a relaxed meditation.

Playing a soundscape through ANC headphones can present a different world acoustically, but so far there has been a lack of research on how augmented footstep sounds affect listeners in such auditively modified worlds in urban areas. We therefore developed *EnvironZen*, a mobile app that augments the footstep sounds to immerse the user into a different world, while visual stimuli of reality are preserved. In order to investigate the influence on walking, we first examined different soundscapes and footstep sounds in an online survey and then tested the most preferred soundscapes in an in-situ field study with 16 participants. Furthermore, we investigated whether augmented footstep sounds are suitable for navigation and observed the influence of sound delays and speedups and the playback of different warning sounds, e. g. when bicycles are approaching. Repeated small interviews were conducted to capture each individual experience and perception of the urban environment and soundscapes.

***Corresponding author: Maximilian Schrapel,** Leibniz University Hannover, HCI Group, Hannover, Germany, e-mail: maximilian.schrapel@hci.uni-hannover.de
**Janko Happe, Michael Rohs,** Leibniz University Hannover, HCI Group, Hannover, Germany, e-mails: janko.happe@gmail.com, michael.rohs@hci.uni-hannover.de

## 2 Related Work

Michael Southworth showed that visually impaired people prefer other urban areas than people without visual impairment, as different characteristics gain importance [90]. From this, Southworth derived the concept of the *sonic city*. R. Murray Schafer established the term of *soundscapes* [82], which is composed of three features: keynote sounds (natural noises like wind or water), signals (acoustic warnings like bells), and soundmarks (individual sounds that identify a specific location). In contrast to an acoustic environment, a soundscape is a deliberate composition of the aforementioned three features by a designer's vision [53]. With the increasing awareness of acoustic environments [53], cities start to actively monitor noise pollution [9], as noise is related to health problems [48, 74], whereas natural sounds have been found to provide health benefits [2, 10, 22].

An audio augmented reality (AAR) is a superimposition of real world soundscapes with virtual sound sources [29]. It embeds computer-generated sounds into the user's acoustic environment [35, 63]. Sonification integrates data into the sounds to convey information [60]. Those acoustically embedded data can, for example, present social media messages [32, 45], location-based information [15], emotions [83], notify specific users [21], help to maintain a steady walking pace [59, 72], or modulate EEG data on soundscapes to assist group walking meditation [27, 28]. One of the first works that investigated auditory displays was Nomadic Radio [81]. Notifications were embedded in the user's environment via wearable loudspeakers and voice recognition was used for input. In a mobile context, AAR is used to promote exploratory and immersive experiences as well as serendipitous discovery of public places [12, 34, 65, 100]. For instance a leisure walk in an AAR park [61, 100] can promote health [73, 77] and support rehabilitation from stress-related mental disorders [23]. For the blind, audio output can describe the natural environment to promote nature connectedness [5, 6]. Playing bird sounds through speakers in noisy urban parks can positively influence an ambience [99] and provide interactivity, e. g. via smart clothing [58]. Following the concept of a *Sound Garden*, anyone can place an auditory virtual park at a specific location to offer listeners a new experience [85, 86] or guide tourists between landmarks [8, 14, 66]. This concept can also be adapted to recreate the soundscape of historical places [88] or to represent historical events [78], e. g. to enhance touristic experiences [26] and museum guides [104] or stories [80]. Thereby, spatial audio can support the design process of a soundscape [46] and 3D sound can contribute to achieving immersion in audio-augmented worlds, e. g. for sound-based games in urban areas [25]. Furthermore, spatial audio can be used for guidance. For example, the direction from which music is played [1, 41], or individual instruments [44] can be used for guidance. To determine the orientation of the user in relation to the played spatial sounds, sensors in the headphones or the smartphone can be used as a virtual directional microphone [42, 43]. Passive or active acoustic isolation, i. e. cocooning, [13] limits the ability to hear oneself (footsteps, voice or breathing) as well as external noise sources. Nowadays, Hearables [31, 76], Earables [55] and headphones with active noise cancellation allow for acoustic transparency to let their listeners stay aware of the environment. McGill et al. envisioned that in the future users will immerse themselves into an acoustically interactive space with personal and private auditory displays [64]. They investigated the impact of acoustic transparency with a focus on external sound sources and found that acoustically transparent and isolated headsets have an effect on the realness and presence of the content.

In virtual realities, footstep sounds enhance the perceived level of presence [56] and also affect walking behavior [47, 84, 101]. In the context of AAR, sonified footsteps offer a certain degree of realism and interactivity with virtual acoustic worlds. Concepts of mounted step sensing and footstep playback on urban running routes aim to provide a more natural running experience [98]. Vibration feedback from shoes was used to change the perceived physical structure of a surface [89, 93]. Different natural surface simulations by footstep sounds were found to influence the biomechanics of walking [95], perceived body weight [92], emotions [62], step length [30], and walking pace [67, 97]. Those effects were also used to manage gait disturbances in Parkinson's disease [79], gait rehabilitation in chronic stroke patients by pitched natural footstep sounds [36], and foot pressure balance for runners [71]. Music that was modified by footsteps showed no perceived effects on user's gait [38]. Using non-naturalistic footstep sounds showed no significant effects on walking patterns [17]. The preferred volume of footstep playback was found not to significantly differ from the volume of the soundscape [96]. Most of the mentioned studies that analyzed augmented footsteps were carried out under controlled lab settings and did not involve users walking in real application scenarios. Hence, the results are often based on small distances. Furthermore, effects due to the measuring equipment, e. g. shoe prototypes, could influence the results.

The existing body of related works shows that soundscapes and footsteps contribute to the immersion in virtual worlds [56], in which footstep sounds also influence walking behavior [47, 62, 67, 84, 92, 95, 97, 101]. There is a positive influence of natural sounds [2, 10, 22, 23] on health, and augmented footsteps sounds may serve as an unobtrusive feedback channel [36, 71, 79]. In addition, soundscapes themselves can provide special cues about the environment [15, 21, 32, 45, 83]. However, so far the main focus of deploying interactive soundscapes has been on city parks [6, 61, 85, 99, 100], guidance with soundscapes [6, 66, 78, 88, 104] and virtual reality with augmented footsteps [47, 56, 84, 101]. Our vision is motivated by noisy and stressful city center environments [7] that can now be masked out by ANC headphones [20, 68, 70]. This might enable the transformation of any acoustic environment into a natural soundscape and provide interactivity in AAR via augmented footstep sounds. Relaxing and exploratory experiences to promote health by walking [73, 77, 100] thus can be implemented at any location and extend the capabilities of sound gardens [85]. The created natural soundscapes can be enhanced by warning sounds that fit into the virtual acoustic environment in order to notify the listeners [21] of hazardous situations, without completely taking the user out of the immersive experience. For instance, approaching vehicles that share their location in real time can communicate with pedestrians in virtual acoustic worlds [24]. However, in previous works, external sounds from the virtual environment were used as cues, while virtual sounds produced by the listeners themselves, e. g., footsteps, were not considered. For example, when accidentally leaving a guided audio path [66], special footstep sounds can notify users that they walk the wrong way, as they would when walking beside a gravel path in the forest. Accelerated or delayed footstep sounds could unobtrusively hint users to hurry up, e. g. to catch a bus, or calm down to be more relaxed [67]. By switching between different soundscapes, users can experience a multifaceted path with a rich variety of impressions [97]. In this work, we explore the vision of augmented interactive soundscapes in city centers with augmented footstep sounds and its influence on listeners. In contrast to [64], we focus on isolated auditory displays and investigate the impact of footstep sounds and their capabilities in audio augmented realities. As a research vehicle we use the *EnvironZen* application. To ensure universal and ubiquitous usage, only smartphone acceleration data are applied to establish real time footstep detection and sonification. This approach does not require instrumenting users with additional equipment that may have an impact on the natural user's gait. The main contributions of this paper include (1) an open-source application for augmenting footstep sounds, (2) an evaluation of individual perceptions of urban stimuli in audio augmented realities, (3) the sonification of augmented footstep sounds to convey navigation cues, and (4) an evaluation of the effects of accelerated and delayed footstep sound feedback.

## 3 EnvironZen Application

*EnvironZen* is designed to transform any acoustic environment into a relaxing soundscape. Based on motion data of the most ubiquitous device – the smartphone – [4] the application aims to recognize and augment footsteps, anywhere and without any additional equipment. The name is composed of *Environment* and *Zen*. Although we do not fully relate our app to the principles of *Zen Buddhism* and walking meditation [39, 50, 51], we aim to encourage users to have a relaxing, stress-relieving walking experience. The idea behind EnvironZen is to motivate a walk, e. g. when going to work or during a lunch break in city centers, where due to the short time of the break no parks are within reach. Since visual stimuli are maintained, the approach is compatible with busy environments, like an urban sidewalk. The auditory footstep augmentation was developed to ideally immerse the user more fully in the auditory world. The state of the Buddha figure in the startup screen of EnvironZen (Figure 1, left) displays the app status. If the Buddha figure is shown with an urban environment in the background, the sound augmentation is off. Swiping up changes the figure: the Buddha now wears

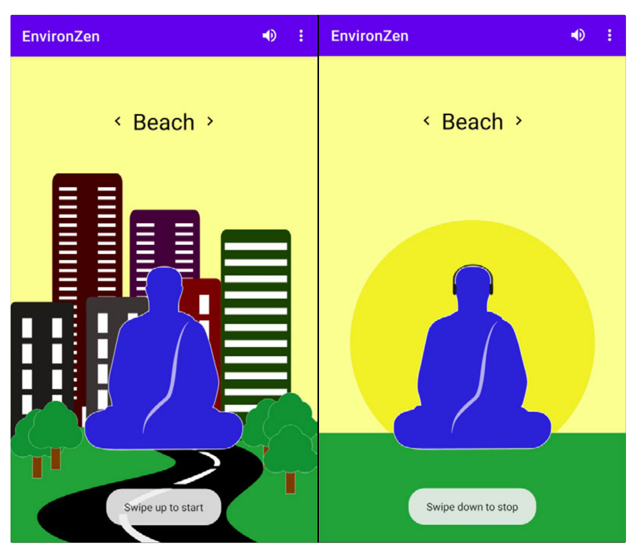



**Figure 1:** Left: Start screen of EnvironZen with inactive soundscape; the urban environment signals the unmodified urban sound environment. Right: Swiping up activates the "Beach" soundscape; swiping left or right selects other soundscapes; swiping down stops the audio augmentation.

headphones and the urban background fades away, being replaced by a circular "halo". The selected soundscape starts playing ("Beach" in Figure 1). Swiping left and right changes the soundscape and footstep sounds. By the background color and the name of the soundscape, users can identify their selection. Footstep sounds are played only when the smartphone is put in the trousers pocket, which is recognized by the proximity sensor of the display. In the menu settings users can also disable the pocket mode. However, EnvironZen's algorithm was optimized for carrying the smartphone in the pocket since the visual perception of the environment is to be preserved. The user can adjust the volume of the footstep sounds and of the soundscape via the speaker button on the top of the screen. Additionally, in the menu users can choose a navigation mode to be guided to a selected location. In this mode, the footstep sound is changed to a wooden step sound when users are walking in the wrong direction. Further, a story mode was implemented that switches randomly between the soundscapes every three minutes, in order to tell the listener a story without any spoken words. The term *story mode* is intended to encourage the user to engage with the virtual acoustic environment. While switching, a ten-second crossfade between the soundscapes and the footstep sounds ensures a pleasant adaptation for the listener. For study purposes, the app can be connected to another smartphone via Bluetooth to start recordings and to activate various study parts. The application was created for Android 6 or higher and is freely available at GitHub[1] along with the audio clips.

## 3.1 Step Detection

Step detection algorithms can analyze motion measurements in either the frequency domain [54] or in the time domain. Time domain approaches include, but are not limited to, thresholding [3], zero-crossing counts [49], and peak detection [103]. Our step detection algorithm is based on the recommendations by Brajdic and Harle [16]. However, we had to extend and reconfigure the proposed windowed peak detection algorithm [16] in order to support real-time step counting. In contrast to Brajdic and Harle, who used a fixed window size of 590 ms for peak detection, we use a window size of 500 ms and change it to 330 ms when the acceleration intensity is above a threshold. We included this adjustment to support activities like short runs, e. g. to catch a bus, for which a fixed window size showed highly inaccurate step counts. Thus, the algorithm covers two states: walking and running.

In Algorithm 1, *MovingAverage(duration)* is an object that computes the moving average over the last *duration* seconds (window size). The *add* method adds another acceleration magnitude value. The *avg* method returns the current moving average. The magnitude is calculated as $accMagn = \sqrt{a_x^2 + a_y^2 + a_z^2} - g$. The earth's gravity ($g = 9.81\,m/s^2$) is subtracted in order to zero the result when no activity is performed. The magnitude is then filtered with a moving average over 310 ms, as proposed by [16]. The parameters $p_1$ and $p_2$ indicate window sizes (in seconds), while $p_3$ to $p_5$ serve as thresholds to detect signal peaks. The parameters have been optimized by an evolutionary algorithm using a Gaussian mutation operator. The window size *minWindow* is initialized with 500 ms and is updated for each measurement depending on the intensity filter. In case the intensity filter exceeds the threshold $p_3$, *minWindow* is set to 330 ms to allow for accurate step detection while running. The assumption here is that the acceleration is higher for running than for walking. If the time since the last recognized step is less than *minWindow*, then no step is reported. A step is recognized when the filtered magnitude *accMagFiltered* is above the threshold *thresholdFilter.avg* multiplied by the parameter $p_4$. Fur-

**Algorithm 1** Step Detection Algorithm.

```
minWindow = 0.500 sec
magFilter = new MovingAverage(0.310 sec)
thresholdFilter = new MovingAverage(p1 sec)
intensityFilter = new MovingAverage(p2 sec)
procedure DETECTSTEP(accMag)
    magFilter.add(accMag)
    accMagFiltered = magFilter.avg
    thresholdFilter.add(accMagFiltered)
    intensityFilter.add(accMagFiltered)
    if intensityFilter.avg > p3 then
        minWindow = 0.330 sec
    else
        minWindow = 0.500 sec
    end if
    if timeSinceLastStep > minWindow then
        if accMagFiltered > thresholdFilter.avg * p4 and accMagFiltered > p5 then
            return true
        end if
    end if
    return false
end procedure
```

1 https://github.com/M-Schrapel/EnvironZen

ther, the filtered magnitude *accMagFiltered* must be above the absolute threshold $p_5$ to detect peaks in the signal.

To optimize this algorithm, two co-authors of this paper (25 years, 75 kg, 1.80 m and 32 years, 97 kg, 1.89 m) recorded accelerometer measurements with a mobile phone (a *Xiaomi 9T* and a *Huawei Mate 20 Pro*, respectively) placed in the front left trousers pocket. They walked for five minutes on sidewalks to record accelerometer data. One developer (32 years) is an experienced runner, recording additional data at an almost constant pace of 10 km/h while running in bear-foot shoes. A total of 230 steps were extracted for the optimization of our algorithm. Samples were discarded for which the magnitude's signal peak did not exactly indicate a heel strike.

The dataset of 230 steps was used to optimize the parameters $p_1$ to $p_5$ with a fitness function that runs the step detection algorithm over the entire dataset until all steps are recognized. The algorithm is provided with the actual number of steps and has to locate them in the dataset. In each iteration, all parameters are mutated with a Gaussian function. The steps are then computed and the new result is compared to the previous result. For this purpose, the *(1 + 1)* evolutionary algorithm [33] selects the previous (parent) or new parameterization (child) with number of recognized steps closest to the real number of steps, i. e., the lowest fitness score, and continues the mutation and calculation until no steps are unrecognized. The resulting parameters are shown in Table 1. The algorithm was subsequently tested by one female and four male volunteers (age: M = 31.6, SD = 13.2 years; height: M = 1.77, SD = 0.1 m; weight: M = 77.6, SD = 8.45 kg). They installed a test application on their own smartphone (two *Huawei One Plus 6*, *Samsung Galaxy S5* and *S10e*, *Google Pixel 4a*) and were introduced to walk for five minutes on a pedestrian walkway with the phone in their left trousers pocket and headphones on. During the test they heard a beep sound every time a step was detected. Although two participants mistakenly put the phone in their right trousers pocket, all testers reported an accurate and instant step detection when their feed touched the ground. To evaluate the performance of the step sensing algorithm, we have built a pair of sensing sandals with attached force sensing resistors on the sole. An Arduino measured both the pressure on the sole and the audio signal amplifier output from a *Zealot H6* Bluetooth headset. The headset is connected to a *Xiaomi 9T* smartphone running our application. We recorded timestamps, when the sole pressure changed and a beep sound is played. On the phone we record the time from a triggered accelerometer sensor event to immediately before a footstep sound is played. One author recorded time measurements from 30 steps

**Table 1:** Algorithm parameterization based on 230 recorded steps from two people during running and walking. The parameters were optimized with an evolutionary algorithm using a Gaussian mutation operator.

| Parameter | Values | |
|---|---|---|
| | Initial | Optimized |
| $p_1$ | 0.2 | 0.629 |
| $p_2$ | 1.5 | 1.365 |
| $p_3$ | 4 | 3.893 |
| $p_4$ | 1 | 1.017 |
| $p_5$ | 0.7 | 0.715 |

walking along a corridor. We measured that the mean time from the foot touching the ground to sound being emitted is 351.9 ms (SD = 104.8 ms), of which on average 300 ms are due to the Bluetooth delay [94]. The average step sound delay of 51.9 ms includes an average of 3.9 ms (SD = 1.4 ms) from the moment the sensor routine detects a step to triggering a sound event on the smartphone.

## 3.2 Soundscapes & Footstep Sounds

We use different royalty-free soundscapes and footstep sounds from the web for our application. Ten different, randomly selected footstep sound variations for each surface ensure that listeners get the feeling of uniqueness for every step they take. Furthermore, the measured duration between two detected heel strikes is used to model the volume of the footstep sound in order to reinforce this impression. The spectrograms shown in Figure 2 visualize some of the implemented soundscapes and their frequency distributions. *Beach* is composed of an ocean wave keynote sound, which is accompanied by seagull cries as soundmarks. In this soundscape, users walk on sand to create the impression of a walk on the beach. The spectrogram shows a relatively uniform frequency distribution, which is primarily defined by the ocean wave peaks. *Creek* is defined only by the keynote of flowing water. Although the spectrum is similar to *Beach*, there are no clear soundmarks here and the listener walks on deep puddles. *Forest* is a rich soundscape with singing birds as soundmarks and a small creek as a keynote sound. The listener walks on dry leaves through this soundscape. The spectrum here is characterized by high-frequency birdsong. Similarly, *Jungle* takes the listener to the soundscape of a rainforest. Besides singing birds, chirping crickets and frogs can be heard. The reverb effect creates a wide and extensive atmosphere, in which the listener walks on the same dry-leaves-sound as before in the forest. The abundance of the

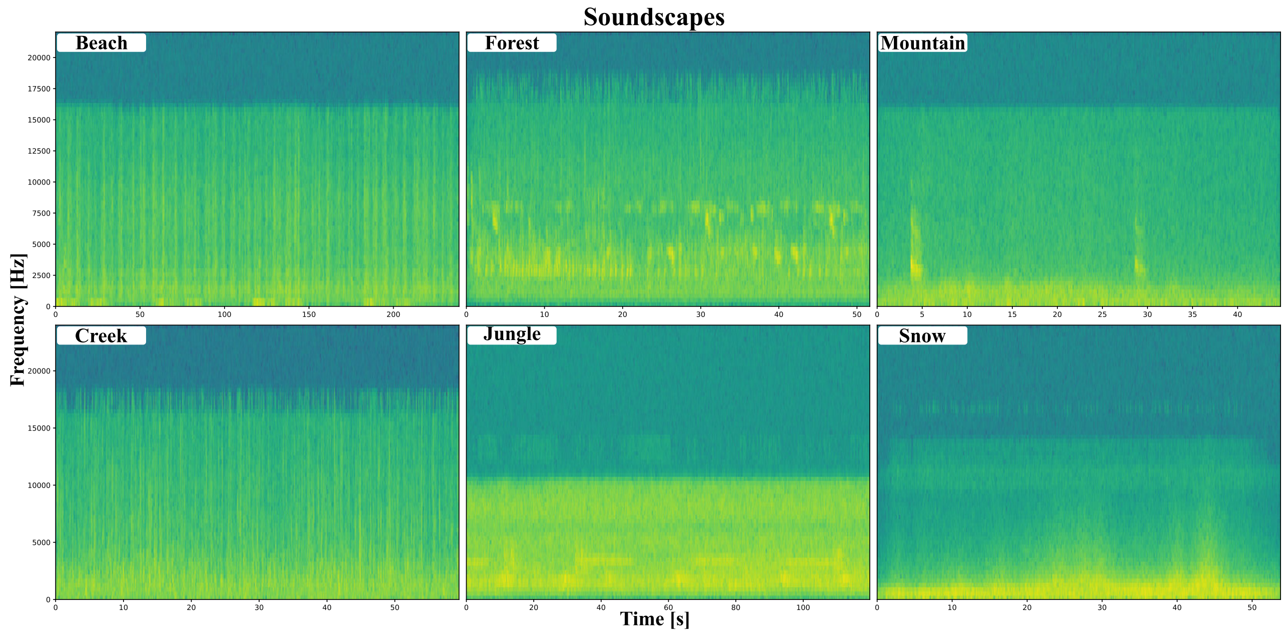



**Figure 2:** Spectrograms of the various soundscapes: During a walk the background sounds are played continuously.

environment can also be seen in the high intensities in the lower frequency range. The *Mountain* soundscape represents an empty place, only defined by blowing wind and a screaming eagle, which is clearly identifiable in the spectrum by the two peaks at 5 and 30 seconds. Here, the listener walks on little stones or gravel whereby the strong wind hints at the location. Likewise, the *Snow* soundscape is defined by cold, blowing wind, which presents the soundscape much more dominantly than the wind in the mountain soundscape. Comparing the spectra of *Snow* and *Mountain* reveals this difference. Here, the listener walks on deep snow through this rough and cold environment. Further, we implemented a *Rain* soundscape, in which continuous rain and a footstep sound of walking in a puddle creates the acoustic environment. The *Musical* soundscape was added as a playful, interactive music environment, in which the act of walking generates music similar to a floor piano. A continuous drum beat is played and to every footstep, a random note of a kalimba is played. In this soundscape, the effect of the footstep volume adjustment based on the duration between two steps becomes particularly useful for creating unique sounds with every step.

## 4 Online Survey

We aimed to evaluate our soundscapes and the footstep sounds in order to find the most preferred and pleasant sounds for a subsequent in-situ field study. In addition, we wanted to investigate which warning signals that match the soundscape can also be used to interpret approaching vehicles, such as scooters or bicycles. Sixteen individuals (four female and twelve male) aged between 22 and 32 years (M = 25.2, SD = 2.3 years) were surveyed. They were asked about their headphone usage and whether they use noise cancelling headphones. Further, we added questions regarding exercises for relaxation purposes with and without digital support through wearables or apps.

After the first set of demographic questions and headphone usage, respondents were introduced to put on their headphones in a quiet environment and to listen carefully to the subsequent audio samples. In the first part, they had to listen and rate the footstep sounds from Section 3.2 on a 5-point scale from very unpleasant to very pleasant. The survey continued with a vision video of our application, in which we introduced the idea of augmented footstep sounds in urban areas. As before, the respondents listened to and rated each soundscape including the footstep sounds. The last part showed a 10-second video of a walking person and an approaching bicycle to demonstrate the idea of soundscape-related warning sounds. The audio track consisted of a soundscape and footstep sounds from Section 3.2. In this video, the approaching cyclist is shown ringing the bell, but the bell sound is replaced by different warning sounds, depending on the soundscape. The warning sounds were a galloping horse for *Beach*, a woodpecker for *Forest*, sleigh bells for *Snow*, the roar sound of a tiger for *Jungle*, and an electric guitar for *Musical*. After each warning sound, we asked the participants to judge how well the sound fits into the scenery and whether the sound is suitable to warn of approaching vehicles. The soundscape *Mountain* was not surveyed on warning sounds, as we here use the bicycle bell as the warning sound in the in-situ study.

## 4.1 Results

Eight participants stated to regularly use headphones while walking and doing outdoor sporting activities. Seven of them use ANC headphones for listening to music in public places, during transportation and at work. Six participants reported to perform meditation exercises like yoga or breathing exercises for relaxation purposes. Four of them support their exercises with mobile applications. Eight stated that they have never used meditation applications for recreational activities. Six of them declared that they had never done meditation exercises before.

Figure 3 depicts the ratings of the sounds presented in the survey. *Forest* was found to be the most relaxing presented soundscape, followed by *Beach*, and *Mountain*. A Friedman test indicated that our soundscapes were rated differently $\mathcal{X}^2(6) = 24.234$, $p < 0.001$. A subsequent post-hoc Tukey test showed that *Forest* was rated significantly higher, while *Musical* was rated lowest among the others. The respondents could not imagine using *Musical* for recreational purposes, but found the *Guitar* to be the most appropriate warning sound, followed by the *Sleigh Bell* and the *Tiger* sound. In the *Forest* soundscape, the *Woodpecker* was appropriate but less dominant as warning sound. The *Horse* at the *Beach* was mentioned to grab attention, but our respondents found the sound inappropriate on the beach. Thus the sound was changed for the in-situ study to a *Ship Horn*, introduced and rated in the opening questionnaire of the in-study. *Leaves* was found to be the most suitable footstep sound for *Forest* and *Jungle*, followed by *Gravel* on the *Mountain* and *Sand* on the *Beach*. *Leaves* was remarked as similar sound to a plastic foil by two respondents. In the *Snow* soundscape participants perceived the footstep sounds on *snow* similar to a wooden floor. From the findings we selected *Forest*, *Beach*, and *Mountain* for the in-situ study. Since the galloping *Horse* in the *Beach* soundscape was rated low, we decided to change that sound to a *Ship Horn* for the main study. Further we asked which other soundscapes they would like to immerse in. Participants mentioned a drip stone cave, the soundscape of the moon, a medieval township, a sunny meadow, New York City at night, and one wanted to have a trampoline sound for the footsteps in a fun park.

# 5 In-Situ Field Study

We conducted a field study to find out how augmented footstep sounds can be used in urban environments. We recruited sixteen participants (three women and 13 men) aged 21 to 58 years (M = 28.9, SD = 9.7 years) for the in-situ study. Seven of these participants (M = 26.9, SD = 2.4 years) also participated in the previous online survey. One day before each volunteer participated in the study they were kindly instructed to fill out an initial questionnaire. We asked the same demographic questions as before about headphone usage and introduced the topic of the survey. In addition, we repeated the warning sound test for two selected soundscapes *Forest* and *Beach*. The intention was to increase awareness of the warning sounds for the fol-

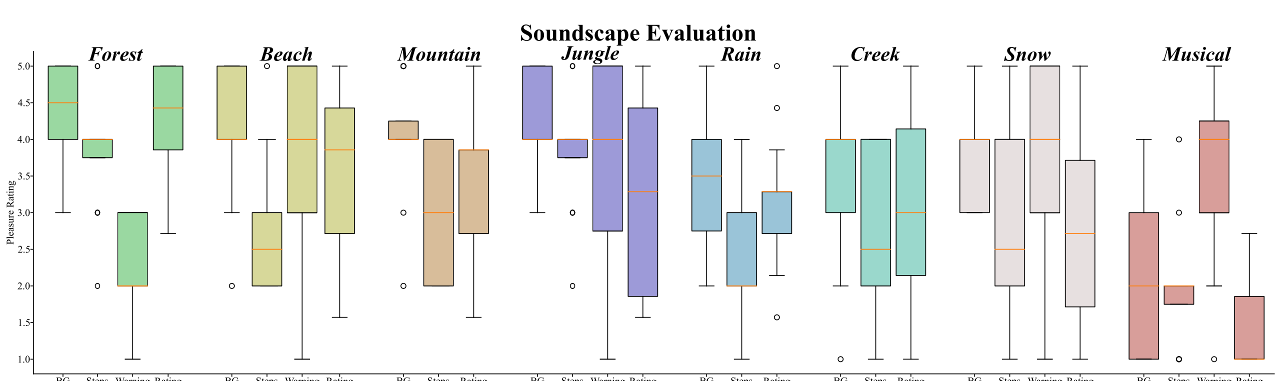



**Figure 3:** Ratings of the implemented soundscapes, sorted from highest to lowest score. A score of 5 presents the highest rating. The pictograms in the bars indicate the footstep sounds. BG is the background environment without any augmentation (just ANC). The first three soundscapes were also evaluated in the in-situ study. The colors of the bars separate the soundscapes.


**Figure 4:** Image at starting point (right) and study route map (left): at route segment *A* participants walked with headphones in ANC mode without any sounds. From *B* to *D* the soundscapes *Beach*, *Forest*, and *Mountain*. Augmented footstep sounds *Sand*, *Leaves*, and *Gravel* were played in the sections marked in green. Segment *E* was used for storytelling with footstep sounds and for playback of warning signals when vehicles are approaching. The pictograms indicate the soundscapes and the yellow marks show construction sites. Segment *F* was used for navigation. *G* and *H* were used for delayed or accelerated footstep sounds using the *Forest* sounds.

lowing study so that the participants are able to interpret these signals in the soundscapes. For the *Beach* soundscape, two sounds were presented. At first the galloping horse sound from the online survey and second a ship horn sound. The latter was chosen for the study as this dominant sound can be more easily associated with a marine environment.

The study took place on weekdays between noon and 5 PM in the city center of Hannover, Germany on a crowed pedestrian sidewalk next to a city street. Two construction sites, cars, restaurants, and stores shaped the lively acoustic environment and thus gave us the opportunity to test our application in the proposed scenario. Figure 4 depicts the study route taken in summer season of 2021 mainly during sunny weather with some wind blowing on several days. At the starting point, the experimenter handed out the *EnvironZen* App installed on a *Huawei Mate 20 Pro* to the participant. During the study the experimenter remotely controlled the app via Bluetooth with a *Huawei OnePlus 3* smartphone. The participant was instructed to comply with the traffic rules and asked put on the *Sony WH-3000-XM3* ANC headphones.

After calibrating the ANC algorithm to the participant's ears, the participant adjusted the volume of the soundscape and the step sounds to their preferred level and put the phone in their left trouser pocket. To make them familiar with ANC, they walked route segment *A* with ANC on, but without any sound augmentation to the end of the first sidewalk. The experimenter followed the participant all the time at a short distance. At the end of the route segment the participants took off the headphones and the experimenter interviewed them about the impression of the noise cancellation. All interviews were audio recorded for later analysis. Afterwards on the route segments *B*, *C*, and *D* the three soundscapes *Beach*, *Forest*, and *Mountain* were played in counterbalanced order. Until the middle of each route (marked in blue in Figure 4) the participants listened to the pure soundscape without footstep augmentation. Then (marked in green in Figure 4), the augmented footstep sounds *Sand*, *Leaves*, or *Gravel* were played depending on the soundscape. At the end of each route segment participants were interviewed about their impression on the temporal accuracy of augmented footstep sounds, whether they experienced immersion into the soundscapes, what they remembered to having heard, and whether and in what way the footstep sounds contributed to the scenery. When participants remarked that the volume of the soundscapes or of the footstep sounds was not appropriate, they had the opportunity to readjust the volume. Furthermore, we recorded GPS and accelerometer data, as well as timestamps, volume, and recognized steps.

Route segment *E* is divided into the three soundscapes *Beach*, *Forest* and *Mountain* with augmented footsteps as shown in Figure 4. This segment was intended to evaluate whether footstep sounds can contribute to more immersive storytelling and how warning sounds are perceived that aim to direct the user's attention to approaching vehicles. The warning sounds are realized with spatial audio in eight compass directions. The intended story was simple: A walk begins on a beach, then leads though a forest, and finally ascends to the top of a mountain. When vehicles like bicycles or e-scooters were approaching, the experimenter sent a command to the participant's smartphone that specified the direction of the vehicle and the soundscape-related warning sound. The warning sounds *Ship horn*, *Woodpecker*, or *Bicycle Bell* were played from the vehicle's direction in relation to the user. This does not provide a controlled study setting, but it does make the participant more aware of approaching vehicles while using ANC headphones for their safety. The sound was linearly cross-faded between two adjacent soundscapes to avoid hard transitions. Before this part began, participants were only told that they were about to hear a story and warning sounds when vehicles are approaching to which they should pay attention to. Afterwards, participants were interviewed about the story and warning sounds they heard.

Route *F* was used to evaluate whether footstep sounds can be employed for navigation purposes. In the *Forest* soundscape participants walked on *Leaves* as long as they walked in the right direction. When they walked on the wrong path, the experimenter remotely changed the footstep sound to *Wooden Planks* with his phone. Before this part began, participants were informed that the system would navigate them a destination unknown to them, which they can identify by a *Gong* sound. Further, they were informed that they would walk in the wrong direction when they hear themselves walking on a wooden surface. When reaching the location, participants were asked how they experienced the feedback.

The aim of the last part on route segments *G* and *H* was to measure the influence of asynchronous footstep sounds. Again the *Forest* soundscape was used together with the *Leaves* footstep sound. The participants were introduced to walk to the end of each route without giving them any information on the intention of this study part. By averaging the time between the first five recognized steps we measured the current walking speed. Based on this calculated average step duration a footstep rhythm continued that, depended on the study part, ether increased or decreased by five milliseconds every two seconds. After 60 seconds, or the initial averaged step duration plus or minus 150 milliseconds, the footstep playback continued with this maximum step duration offset. Each participant either began by hearing increasing or decreasing footstep sound periods on route segment *G* and then continued with the other method on route segment *H*. After each segment, participants were interviewed if they noticed anything. Afterwards, participants were handed out a final questionnaire for gathering feedback about this study and the presented technology. We asked which study part they liked most, which soundscape they preferred and why, how accurate they found the timing of the step sounds to be, whether step sounds provided a more immersive and relaxing experience, as well as in which context they would use such audio augmentations. Furthermore, the participants were given the opportunity to give us free-form feedback.

# 6 Evaluation

In the following, the results of the in-situ field study are presented.

## 6.1 Augmented Footsteps vs. No Augmentation

Initially, the participants walked with active noise cancelling to familiarize themselves with the technology. In the subsequent interview eleven participants mentioned that it was unusual but pleasant to block the environmental noise. One noted that he never before realized how noisy the environment actually is. Two remarked that the headphones amplified their own footstep sounds while walking. In the following, results of our interviews from the counterbalanced soundscapes are presented.

In the *Beach* soundscape, participants mostly mentioned the ocean wave sound first as a dominant keynote sound. One participant thought he mistook the sound of waves for approaching cars, and another noted that he had to look around to see where the water was coming from. P2 stated *"It was confusing, it felt like I'm walking on a beach but I couldn't see it."*. Similarly P8 said *"It was surreal to walk on the lively street but hear the soundscape of a beach."*. Interestingly, some participants incorporated their visual impressions into the soundscape and felt even more immersed in a beach environment. P1 and P11 remarked that they had the impression of walking on a beach promenade, because of the restaurants nearby. P10, P12, and P14 had been on a beach vacation shortly before and felt taken back by the soundscape. With footstep augmentation, P6 noted *"With the step sound it was different, I was more focused on the acoustic environment, it was more intense."*. However, P16 perceived, in contrast to the other participants, the steps as not being in sync with his footsteps and felt like someone is walking behind, which he found confusing. P10 remarked that augmented footsteps suppressed the sound of his own footsteps, which made the soundscape more immersive. P11 mentioned that the augmented footsteps reinforced the soundscape and gave a more intense feeling, making him focus more on his own footsteps and breathing. P8, P9, P12, P14 and P16 found the footstep sounds slightly too loud, which made it more difficult for them to immerse themselves in the environment. P12 mentioned that he was able to get used to the volume after taking a few steps. The participants were therefore given the opportunity to readjust the footstep volume for the following parts.

When the experimenter asked after the *Forest* soundscape what kind of soundscape they heard, seven participants mentioned the *Leaves* step sound as a remarkable environmental feature. P15 rather mentioned dry grass and perceived the environment, like P10 and P16, in the very deep forest. P1 found the environment with step sounds more immersive but remarked *"The footstep sounds came a bit too early. But when I began to imagine dry leaves on the ground making sounds before the feet touches the ground, I felt even more immersed in the forest."*. P9 mentioned that the footstep sound matched his gait making the soundscape more immersive. For P5 it was challenging to concentrate on the soundscape due to the wind during the study. A misaligned footstep sound playback was perceived as disturbing. P6 repeated the perceived effect of focusing as in the *Beach* soundscape. For P4, the footstep sound increased the presence and the previously decreased footstep volume made it more realistic. P13 remarked that *"the step sound was too beautiful to be realistic."*. For P10, P14 and P15 the step volume was perceived as too loud and had to be adjusted. P12 stated that he immersed more in the soundscape by looking at the trees on the side of the road. Similarly, P3 reported that he focused more on his visual perception.

The soundscape *Mountain* was perceived very differently among the participants. P3 imagined to walk in a forest on puddles which was less pleasant to him. P7 interpreted wind sounds as ocean waves and found the footstep volume too loud, which was readjusted in consequence. P10 found himself in a primeval forest with a screaming parrot. During the interview, when mentioning the walked stony path, he noticed that the soundscape fits better to a mountain after all. P11 could not distinguish whether the wind noise was coming from cars on the road or the soundscape. P14 found the footstep sounds less synchronized with his gait. In contrast, P1 liked the quiet sounds. Although he perceived the soundscape as less immersive, he favored the footstep sound. Similarly, P15 mentioned *"I could almost feel the surface structure of the ground."*. P2 favoured the empty, quiet soundscape. The footsteps on the dry path and the wind were perceived as very relaxing. Readjusting the footstep sounds was remarked as more pleasant for P13. P12 associated the footsteps with a swamp. He remarked that in his bear-foot shoes he felt on the side way with the footstep sounds like he would *"step into something."*. Thus, he decided to shortly walk on a small grass field beside the road which was more pleasant to him. P6 used the quiet soundscape to focus on his gait while walking. Most interestingly P16 smiled and said *"I felt taken back to my childhood. The soundscape matched the place where I grew up. I walked again on the stony path next to corn fields and a forest on the other side."*. Further he continued *"Without the footstep sounds I couldn't have immersed myself back."*.

## 6.2 Storytelling and Warning Sounds

On the story path back to the starting point warning sounds were interpreted correctly, but none of the participants really integrated the approaching vehicles in their imagined story. P1 found the composition made of *Beach, Forest, Mountain* like a short walk through Japan. Beginning from the sea he walked up to the mountains. He remarked the *Ship Horn* as very appropriate to catch his attention. P2 interpreted the ship as a train and could not imagine where it was coming from in the *Beach* soundscape story. The windy weather on this day contributed to immersing into the entire story. Although the *Bicycle Bell* was easy to interpret in the *Mountain* soundscape, the participant was less aware of the traffic and instead felt more relaxed. P2 also stated *"I would like to have control when warning sounds are played. When a bicycle was approaching in front of me I could already see it."*. P3 has perceived the story of a plane crash on a lonely island. He remarked that the *Woodpecker* sound was less dominant and not appropriate for warnings. He watched the traffic but mentioned that he could also immerse into the story. Similarly, P4 stranded on a lonely island. P5 made a travel around Europe starting from the Northsea to the Alps. P6 wandered to a river source and completely forgot the study situation. The crossfade of the soundscapes was remarked as *"extremely convincing."*. P7 felt teleported through different worlds and almost forgot the city environment. The *Woodpecker* warning sound was remarked as *"surprisingly suitable and fitting into the soundscape."*. P8 found the *Bicycle Bell* funny and surprising. It fitted well into the overall reality. P9 made a relaxing day on the beach and went back home in the mountains. Similarly, P10 imagined a tropical island and returned to the resort. He even felt that he could concentrate better on the traffic. P11 felt remembered of a full vacation. P12 made a walk to mountains but had the feeling he could concentrate less on the traffic. P13, P14 and P16 made similar walks. P15 stranded after shipwreck on a lonely island in search for food. The warning sounds did not match his story, which made him feel taken out of his story.

## 6.3 Navigation

Participants were asked how synchronous they perceive the footstep sounds now after using the technology for a while. In 8 cases (P1, P2, P6, P12, P13, P14, P15 and P16) participants remarked that the step sound is now more synchronous as it was in the beginning. P5 was the only one who still perceived the footstep sounds as asynchronous with slight improvements. P3, P7 and P14 were impressed about the synchronous footstep feedback. P8 commented that the synchronicity of the footsteps depends on the sound and its dynamics. P16 remarked that he had to get used to the footstep sounds in the first five minutes of the study. Afterwards, he perceived the sounds and his footsteps as part of the soundscapes. The intended navigation path was walked by two people without taking the wrong path. Eleven participants stated that they instantly understood the footstep feedback on the wrong path and changed the direction immediately. P12 commented that this kind of feedback has a *"binary fashion"* meaning that when there are more than two possible directions, he had to try the other paths. P13 and P16 also mentioned that. P11 would like to hear the feedback left or right to identify the right way. When P2 tried the directions she felt watched by strangers.

## 6.4 Asynchronous Footstep Playback

In 14 cases, people noticed the faster footstep playback. P5 found the step sounds to be asynchronous over the entire study and couldn't identify any changes. P8 could also not perceive any differences although he previously found the step sound to be mostly synchronous. Five participants perceived the faster feedback as stressing, five felt pushed by the footstep sounds and four remarked the feedback as unpleasant. P14 mentioned a feeling of walking beside a person or dog. P16 noticed that he started to walk faster and then decided to slow down his pace. P15 walked his own pace but felt like walking faster.

In contrast, nine participants found no difference when the footstep sounds were played slower. Again, P5 didn't perceived any differences. P10 and P11 thought the slower pace might be an imagination. Five participants (P1, P2, P3, P4, P11) found the slower pace to be relaxing or pleasant. Four participants (P2, P9, P11, P13) reported an urge to slow their pace. P14 again reported the feeling of walking beside another person. He also perceived the walked surface as *muddy*. P2 was persuaded by the footstep sounds and stated afterwards: *"When I noticed the slower footstep rhythm I realized that I really walk too fast and should slow down myself to be more relaxed."*.

In Figure 5 the influence of the playback is visualized. A Mann-Kendall test was performed for each study part. We could not identify a significant trend when footsteps were played faster ($p = 0.79$, $z = 0.27$, $\tau = -0.04$, $s =$

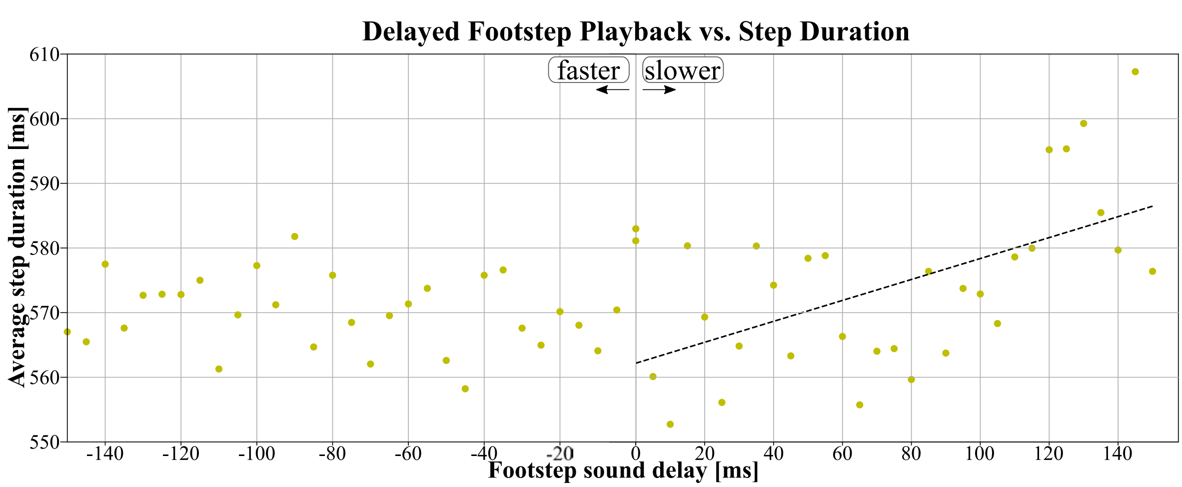



**Figure 5:** Footstep playback delay vs. step duration: A negative delay corresponds to a faster walking speed. The dashed line visualises the significant trend found by a Mann-Kendall test. By decreasing the step duration we could not identify trends.

−17.0). When footstep sounds were played slower, a Mann-Kendall Test identified a significant increasing trend of the step durations ($p = 0.0072$, $z = 2.69$, $\tau = 0.34$, $s = 0.0007$). From the Figure 5, the trends can also be identified.

## 7 Discussion

Soundscapes do not necessarily have to compose a rich sound environment to be perceived as pleasant. In our online survey the *Jungle* soundscape, which has the highest variety of sounds and reverb effects, was not rated higher than the *Forest* soundscape. We nevertheless assume that the preferences are user-dependent. However, the participants of the in-situ study stated to incorporate their visual impressions into the played soundscapes. For example, trees on the side of the road were woven into the acoustic forest and restaurants were integrated into an imagined walk down a beach promenade. To improve the overall experience, the selection of a soundscape could be based on visual impressions of the walked paths. Such “auditory anchors” may be stores and restaurants along the way or particular landmarks, which could be retrieved from map data.

From the moment at which step sounds were played, the environment was experienced as richer and more compelling. Augmented footstep sounds added a new quality to the experience. For instance, participants perceived the acoustic forest as more realistic when they heard the sound of their own footsteps on dry leaves; and P16 imagined to be back to a specific place of his childhood in the *Mountain* soundscape. In our interviews participants mentioned the word *focused* (on steps and breath) when augmented step sounds were played. Therefore, we wanted to find out whether the independent variables soundscapes and footstep sounds have effects on the users' step duration in the first study part. A Shapiro-Wilk test showed no normal distribution of the step duration with augmented footsteps in the soundscapes *Beach* and *Mountain*. Therefore, a Friedman test was performed that showed no significant differences between the three soundscapes with and without footstep augmentation regarding the step duration ($F(6, 16) = 5.86, p = 0.32 > 0.05$). One reason for this outcome may be that the participants had to remain aware of their surroundings, e. g. to circumvent other pedestrians. However, this contradicts findings from related work in lab studies in which the step length and walking pace were both affected by different footstep sounds [30, 97]. In addition to external influences, the footstep sound volume was sometimes perceived to be too loud after initial attempts, although all sounds were normalized and related work found no significant difference in the preferred volume of soundscapes and footsteps [96]. Because a Shapiro-Wilk test confirmed a normal distribution of the volumes of soundscapes and footstep sounds in each study part, we performed a t-test for each pair of volumes. We found that each pair of soundscape volume and footstep volume was significantly different ($t(2, 16) = 13.95, p < 0.001$). One-way ANOVAs were conducted to analyze the volume change of the soundscape and footstep sounds over the study tasks. The results revealed that once footstep sound volume is set, there is no further significant different change of the volume ($F(8, 16) = 1.60, p = 0.14 > 0.05$). The same applies to the soundscape volume ($F(8, 16) = 0.04, p = 0.99 > 0.05$). We therefore conclude, in contrast to [96], that the footstep volume should be adjusted below the soundscape volume. On average, the footstep volume ($\bar{vol} = 42.52\,\%$; $SD = 14.82\,\%$) was set 17.4 % below the soundscape volume ($\bar{vol} = 59.95\,\%$; $SD = 12.22\,\%$) for a pleasant experience. Although no other significant differences were found, a Mann-Kendall test found a significant decreasing trend for the footstep volume change over the study parts ($p = 0.006 < 0.05$, $z = 2.75$, $\tau = -0.16$, $s = -1332.0$) while no trend on the soundscape volume change could be identified ($p = 0.65 > 0.05$, $z = 0.45$, $\tau = -0.16$, $s = -0.03$). Figure 6 depicts the volume change of soundscapes and footsteps over the study.

Our participants reported that they perceived the footstep playback as more synchronous with their actual footsteps after having used the augmentation for a while. This is most likely not only due to a learning effect, but also depends on the volume of the step sounds. An asynchronous step feedback was also experienced as walking next to a person. According to P16, it took him 5 minutes to get used to the feedback and then he perceived the augmented footstep sounds as part of the soundscape. We conclude that a highly accurate footstep synchronisation is not necessary in the context of AAR. In the closing questionnaire the synchronicity of the footstep sounds on a 5-point scale from 1 (fully asynchronous) to 5 (fully synchronous) has been evaluated (M = 4.13, SD = 0.93). Ten participants of the in-situ study rated the synchronicity as 4 (synchronous). Five participants stated full synchronicity (P3, P6, P7, P11, P16) and only P5 found the sounds to be fully asynchronous. Also for P5, a delay of the step sounds led to an increase in the step duration, according to a Mann-Kendall test ($p = 0.04 < 0.05$, $z = 1.97$, $\tau = 0.51$, $s = 23.0$), although he noticed no change in the rate of the footstep playback. Thus, we infer that delayed footstep playback may serve as an unobtrusive feedback method to slow pedestrians down. In contrast, a faster playback was perceived as stressful

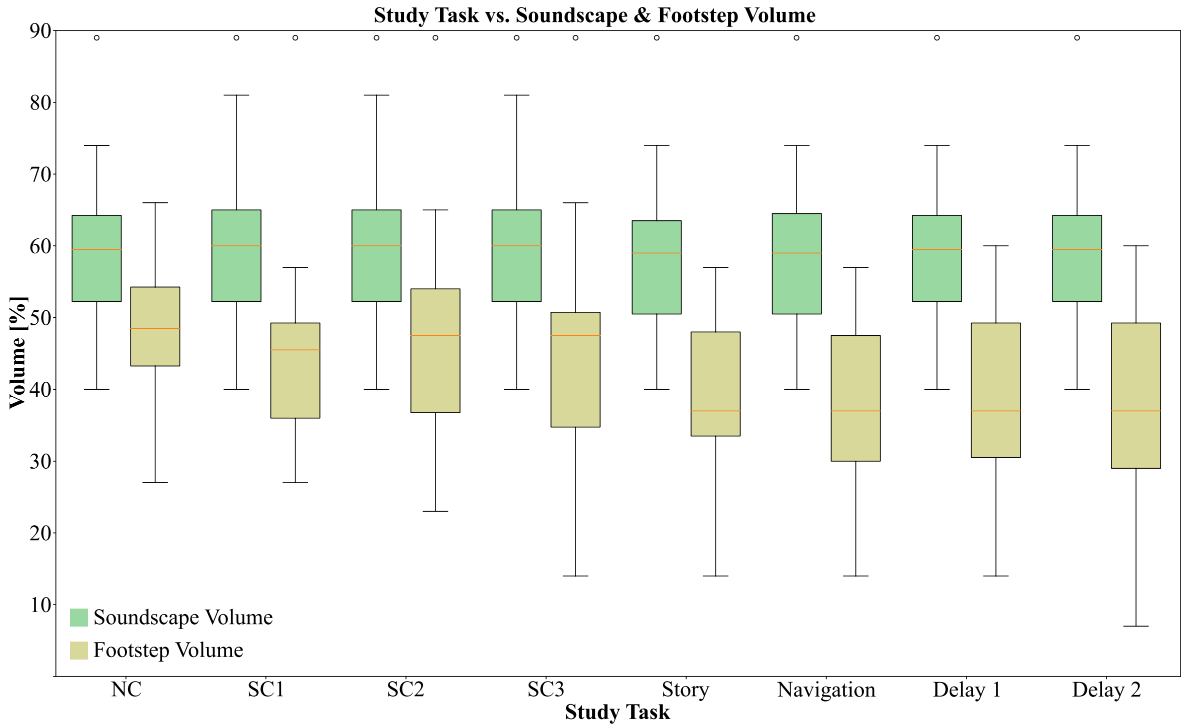



**Figure 6:** Volume adjustment over the course of the study: While both sounds were adjusted equally in the beginning, participants repeatedly lowered the footstep volume (in yellow) from an average of 60 % to 42.5 %.

and pushing, which might be helpful in the context of running with the aim of holding a certain pace or to catch a bus during a navigation tasks. We also evaluated whether the participants were able to immerse themselves in the soundscapes on a 5-point scale from 1 (no immersion) to 5 (full acoustic immersion) ($M = 3.94$, $SD = 0.66$). Three participants totally agreed to have felt immersed into another world (P2, P6, P11), nine agreed and only four were neutral (P4, P8, P10, P13). Here we conclude that footstep augmentation contributes to immersing into audio-augmented realities. Commonly used questionnaires that measure immersion or presence [102] were not applied, as augmented soundscapes fully preserve all visual stimuli of reality and thus several items on standard questionnaires do not apply [64]. Furthermore, the urban environment in the in-situ study presented an uncontrolled setting, leading to confounds when trying to measure immersion. We found that the participants combined external visual impressions and even wind sounds and their tactile and haptic perception into a unified experience. Therefore, we felt that interviews would be a more appropriate method to assess personal experiences of augmented soundscapes in public places [18].

By telling the participants that they will hear an auditory story, they were easily able to make up story lines by themselves. Examples of stories they imagined were: a walk in Japan, being on a vacation, stranding on an lonely island, or even going back to their childhood. The auditory soundscape inspired the imagination of fictional stories or the evocation of memories, whereby the sounds of footsteps influence the imagination. Interestingly, none of our participants talked about a story in the third person. They were always referring to themselves as the main actor in the story, which suggests at least some degree of immersion into the played soundscapes. Mostly the participants imagined being surrounded by a natural environment without other people, which indicates that the participants mentally blocked out the crowded urban environment. Man-made warning sounds, like the bicycle bell and the ship horn, were perceived as surprising and could not be woven into their stories. Instead, those sounds brought the participants out of the story to focus more on the surrounding traffic. The woodpecker signal sound was a better fit for the soundscape, but it may lead to important warning signals being ignored. Hence, it seems more suitable for signalling a change of direction in a navigation task rather than as a warning sound. The binary fashion of the footstep feedback during the navigation task (standard sound when on the right track, wood sound when on the wrong track) was found to be sufficient to signal wrong walking directions. Although this simple navigation feedback was perceived as easy and intuitively understandable, it does not provide sufficiently strong cues to replace spatial audio navigation methods in terms of efficiency [1, 41]. Thus, it should rather be used to deliver a stop signal, e. g. at a red traffic light, or to indicate when walking in the wrong direction. In terms of walking meditation, an audio augmented reality using footstep feedback can offer a new dimension of meditation. Individuals reported feeling more focused on their own steps and breath with footstep feedback. In addition, an AAR should not be considered as a replacement for walking meditation in urban soundscapes but as one potential element of a meditation method [40]. Further research is required to compare the effect of our approach on walking meditation and the effect of AAR without augmented footstep sounds.

In comparison to McGill et al. [64], our research confirms some of their results. Auditory displays can weave into the overall perceived environment. However, the visual influences and footstep sounds have an additional impact on how real an audio augmented reality is perceived. Embedded warning sounds can restore a certain perceived sense of safety with ANC headphones to make the user aware of their surroundings. Likewise, footstep sounds can be used to intuitively, instantly and partly unconsciously convey simple cues. This makes footstep sounds suitable to not overload an auditory display. However, our results show that these sounds require a short period of familiarization to be perceived as part of the overall reality. In addition, the footstep volume needs to be adjusted below the background soundscape volume to provide a more realistic experience, and faster step feedback should be avoided as it was perceived as stressful.

# 8 Conclusion

This paper presented the results of an online survey and a subsequent in-situ field study on augmenting footstep sounds in audio-augmented realities. The field study took place in an urban environment. The soundscapes (*Forest*, *Beach*, and *Mountain*) and footstep sounds (*Leaves*, *Sand*, and *Gravel*) that received the highest ratings in the online survey were investigated in detail regarding their effects on the experience of pedestrians on a sidewalk. Audio was played with noise-cancelling headphones and footsteps were detected with the accelerometer of a mobile phone worn in the trouser pocket.

We found that augmented footstep sounds contribute to the experience of presence in acoustic environments. Augmented footstep sounds create a link between acting in the physical environment – performing the footsteps, experiencing the haptic feedback, and visually perceiving one's movement in the urban space – and in the virtual environment – hearing the virtual footstep sounds in the context of the soundscape. So users act simultaneously in the physical and the virtual world. It seems that the sounds of the preferred soundscapes and of the augmented footsteps are sufficiently indeterminate that users are able to integrate them with the other stimuli to a coherent whole. Users seem to be able to concurrently allocate attention to the auditory stimuli from the soundscape, the virtual sounds created by their footsteps, and the visual stimuli from the physical environment. Here, the imaginations were very different from person to person. When soundscapes are sequentially played, listeners are able to connect them to form stories in urban environments, in which they themselves act as the main characters. Soundscapes can also evoke memories like vacations and these memories or imaginations may even be the main factor determining the level of immersion in quiet, acoustic compositions.

Adjusting the volume of footstep sounds was found to be necessary for creating pleasurable, individual experiences. Slightly delayed footstep playback showed a significant impact on the step duration, making it possible to slow people down, while faster footstep playback showed no impact on walking speed but led to users feeling stressed and rushed. Slower footstep sounds were perceived as relaxing and calming. Using footstep sounds for navigation purposes can be intuitively understood, but is limited in effectively conveying a sense of direction. However, changing the footstep sound can be used to indicate the wrong direction, red traffic lights, or the right way with only two choices. Augmenting further sounds with spatial audio in soundscapes, like approaching vehicles, can be easily understood but should be introduced beforehand for not being misunderstood. One advantage of audio-only augmentation for public spaces is that visual stimuli are fully preserved, so that possible harms only have to be augmented if out of sight of the user.

The underlying algorithm for real-time footstep playback was found to be mostly accurate in the study, but could be adapted to each user to improve the reliability of the acoustic immersion. One advantage of the approach is that it can be implemented and used on any smartphone without additional equipment. *EnvironZen* is freely available at GitHub[2] and open-sourced for future research. As people need to continue to pay attention to events in the physical environment, safety is an important issue to be further explored in the future.

2 https://github.com/M-Schrapel/EnvironZen

# References


[1] Robert Albrecht, Riitta Väänänen, and Tapio Lokki. 2016. Guided by music: pedestrian and cyclist navigation with route and beacon guidance. *Personal and Ubiquitous Computing* 20 (02 2016). https://doi.org/10.1007/s00779-016-0906-z.

[2] Jesper J. Alvarsson, Stefan Wiens, and Mats E. Nilsson. 2010. Stress recovery during exposure to nature sound and environmental noise. *International journal of environmental research and public health* 7, 3 (2010), 1036–1046.

[3] Moustafa Alzantot and Moustafa Youssef. 2012. UPTIME: Ubiquitous pedestrian tracking using mobile phones. In *2012 IEEE Wireless Communications and Networking Conference (WCNC)*. 3204–3209. https://doi.org/10.1109/WCNC.2012.6214359.

[4] Rafael Ballagas, Jan Borchers, Michael Rohs, and Jennifer G. Sheridan. 2006. The smart phone: a ubiquitous input device. *IEEE pervasive computing* 5, 1 (2006), 70–77.

[5] Maryam Bandukda and Catherine Holloway. 2020. Audio AR to Support Nature Connectedness in People with Visual Disabilities *(UbiComp-ISWC '20)*. Association for Computing Machinery, New York, NY, USA, 204–207. https://doi.org/10.1145/3410530.3414332.

[6] Maryam Bandukda, Catherine Holloway, Aneesha Singh, and Nadia Berthouze. 2020. PLACES: A Framework for Supporting Blind and Partially Sighted People in Outdoor Leisure Activities. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20)*. Association for Computing Machinery, New York, NY, USA, Article 29, 13 pages. https://doi.org/10.1145/3373625.3417001.

[7] Leah Barclay. 2017. Augmenting Urban Space with environmental Soundscapes and mobile technologies. *Soundscape. Sounds Emergent: Diverse Ecologies Part II* 16 (2017), 20–34.

[8] Benjamin B. Bederson. 1995. Audio augmented reality: a prototype automated tour guide. In *Conference companion on Human factors in computing systems*. 210–211.

[9] Juan P. Bello, Claudio Silva, Oded Nov, R. Luke Dubois, Anish Arora, Justin Salamon, Charles Mydlarz, and Harish Doraiswamy. 2019. SONYC: A System for Monitoring, Analyzing, and Mitigating Urban Noise Pollution. *Commun. ACM* 62, 2 (Jan. 2019), 68–77. https://doi.org/10.1145/3224204.

[10] Jacob A. Benfield, B. Derrick Taff, Peter Newman, and Joshua Smyth. 2014. Natural sound facilitates mood recovery. *Ecopsychology* 6, 3 (2014), 183–188.

[11] Birgitta Berglund, Thomas Lindvall, Dietrich H. Schwela, World Health Organization, et al. 1999. Guidelines for community noise. (1999).

[12] Liam Betsworth, Nitendra Rajput, Saurabh Srivastava, and Matt Jones. 2013. Audvert: Using spatial audio to gain a sense of place. In *IFIP Conference on Human-Computer Interaction*. Springer, 455–462.

[13] Karin Bijsterveld. 2010. Acoustic cocooning: How the car became a place to unwind. *The Senses and Society* 5, 2 (2010), 189–211.

[14] Costas Boletsis and Dimitra Chasanidou. 2018. Smart tourism in cities: Exploring urban destinations with audio augmented reality. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*. 515–521.

[15] B. Boren, M. Musick, J. Grossman, and Agnieszka Roginska. 2014. I Hear NY4D: Hybrid Acoustic and Augmented Auditory Display for Urban Soundscapes. In *International Conference on Auditory Display*.

[16] Agata Brajdic and R. Harle. 2013. Walk detection and step counting on unconstrained smartphones. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing* (2013).

[17] Roberto Bresin, Anna Dewitt, Stefano Papetti, Marco Civolani, and Federico Fontana. 2010. Expressive sonification of footstep sounds. *Proceedings of the Interaction Sonification workshop (ISon)*.

[18] Bennett M. Brooks, Brigitte Schulte-Fortkamp, Kay S. Voigt, and Alex U. Case. 2014. Exploring our sonic environment through soundscape research & theory. *Acoustics today* 10, 1 (2014), 30–40.

[19] Henrik Brumm. 2004. The impact of environmental noise on song amplitude in a territorial bird. *Journal of Animal Ecology* 73, 3 (2004), 434–440. https://doi.org/10.1111/j.0021-8790.2004.00814.x arXiv:https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/j.0021-8790.2004.00814.x.

[20] Samantha Butterbaugh and Monica Gordon-Pershey. 2019. The Usefulness of Active Noise Cancelling Headphones when Listening to Music in the Presence of Ambient Noise. *eHearsay* 9, 2 (2019), 17.

[21] Andreas Butz and Ralf Jung. 2005. Seamless User Notification in Ambient Soundscapes. In *Proceedings of the 10th International Conference on Intelligent User Interfaces (IUI '05)*. Association for Computing Machinery, New York, NY, USA, 320–322. https://doi.org/10.1145/1040830.1040914.

[22] Rachel T. Buxton, Amber L. Pearson, Claudia Allou, Kurt Fristrup, and George Wittemyer. 2021. A synthesis of health benefits of natural sounds and their distribution in national parks. *Proceedings of the National Academy of Sciences* 118, 14 (2021). https://doi.org/10.1073/pnas.2013097118 arXiv:https://www.pnas.org/content/118/14/e2013097118.full.pdf.

[23] Gunnar Cerwén, Eja Pedersen, and Anna-María Pálsdóttir. 2016. The role of soundscape in nature-based rehabilitation: A patient perspective. *International journal of environmental research and public health* 13, 12 (2016), 1229.

[24] Chia-Ming Chang, Koki Toda, Daisuke Sakamoto, and Takeo Igarashi. 2017. Eyes on a Car: An Interface Design for Communication between an Autonomous Car and a Pedestrian *(AutomotiveUI '17)*. Association for Computing Machinery, New York, NY, USA, 65–73. https://doi.org/10.1145/3122986.3122989.

[25] Thomas Chatzidimitris, Damianos Gavalas, and Despina Michael. 2016. SoundPacman: Audio augmented reality in location-based games. In *2016 18th Mediterranean Electrotechnical Conference (MELECON)*. 1–6. https://doi.org/10.1109/MELCON.2016.7495414.

[26] Uipil Chong and Shokhzod Alimardanov. 2021. Audio Augmented Reality Using Unity for Marine Tourism. In *Intelligent Human Computer Interaction*, Madhusudan Singh, Dae-Ki Kang, Jong-Ha Lee, Uma Shanker Tiwary, Dhananjay Singh, and Wan-Young Chung (Eds.). Springer International Publishing, Cham, 303–311.

[27] Karen Cochrane, Lian Loke, Matthew Leete, Andrew Campbell, and Naseem Ahmadpour. 2021. Understanding the First Person Experience of Walking Mindfulness Meditation Facilitated by EEG Modulated Interactive Soundscape. In *Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '21)*. Association for Computing Machinery, New York, NY, USA, Article 18, 17 pages. https://doi.org/10.1145/3430524.3440637.

[28] Karen Anne Cochrane, Lian Loke, Andrew Campbell, Matthew Leete, and Naseem Ahmadpour. 2020. An Interactive Soundscape to Assist Group Walking Mindfulness Meditation. In *Proceedings of the 7th International Conference on Movement and Computing (MOCO '20)*. Association for Computing Machinery, New York, NY, USA, Article 21, 3 pages. https://doi.org/10.1145/3401956.3404240.

[29] M. Cohen, S. Aoki, and N. Koizumi. 1993. Augmented audio reality: telepresence/VR hybrid acoustic environments. In *Proceedings of 1993 2nd IEEE International Workshop on Robot and Human Communication*. 361–364. https://doi.org/10.1109/ROMAN.1993.367692.

[30] Tara Cornwell, Jane Woodward, Mengnan Wu, Brennan Jackson, Pamela Souza, Jonathan Siegel, Sumitrajit Dhar, Keith E. Gordon, et al.2020. Walking with Ears: Altered Auditory Feedback Impacts Gait Step Length in Older Adults. *Frontiers in Sports and Active Living* 2 (2020), 38.

[31] Poppy Crum. 2019. Hearables: Here come the: Technology tucked inside your ears will augment your daily life. *IEEE Spectrum* 56, 5 (2019), 38–43.

[32] Tilman Dingler and Stephen Brewster. 2010. AudioFeeds: A Mobile Auditory Application for Monitoring Online Activities *(MM '10)*. Association for Computing Machinery, New York, NY, USA, 1067–1070. https://doi.org/10.1145/1873951.1874151.

[33] Stefan Droste, Thomas Jansen, and Ingo Wegener. 2002. On the analysis of the (1 + 1) evolutionary algorithm. *Theoretical Computer Science* 276, 1–2 (2002), 51–81.

[34] G. Eckel. 2001. Immersive audio-augmented environments: the LISTEN project. In *Proceedings Fifth International Conference on Information Visualisation*. 571–573. https://doi.org/10.1109/IV.2001.942112.

[35] H. Gamper. 2014. Enabling technologies for audio augmented reality systems.

[36] Alba Gomez-Andres, Jennifer Grau-Sánchez, Esther Duarte, Antoni Rodriguez-Fornells, and Ana Tajadura-Jiménez. 2020. Enriching footsteps sounds in gait rehabilitation in chronic stroke patients: a pilot study. *Annals of the New York Academy of Sciences* 1467, 1 (2020), 48–59. https://doi.org/10.1111/nyas.14276 arXiv:https://nyaspubs.onlinelibrary.wiley.com/doi/pdf/10.1111/nyas.14276.

[37] Gabriel Haas, Evgeny Stemasov, and Enrico Rukzio. 2018. Can't You Hear Me? Investigating Personal Soundscape Curation. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia (MUM 2018)*. Association for Computing Machinery, New York, NY, USA, 59–69. https://doi.org/10.1145/3282894.3282897.

[38] Nassrin Hajinejad, Barbara Grüter, Licínio Roque, and Simon Bogutzky. 2016. GangKlang: Facilitating a movement-oriented walking experience through sonic interaction. In *Proceedings of the Audio Mostly 2016*. 202–208.

[39] Thick Nhat Hanh. 1985. A guide to walking meditation. *Fellowship* 51, 10–11 (1985), 23.

[40] Thich Nhat Hanh. 2011. *The long road turns to joy: A guide to walking meditation*. Parallax Press.

[41] Florian Heller, Jelco Adamczyk, and Kris Luyten. 2020. Attracktion: Field Evaluation of Multi-Track Audio as Unobtrusive Cues for Pedestrian Navigation. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '20)*. Association for Computing Machinery, New York, NY, USA, Article 21, 7 pages. https://doi.org/10.1145/3379503.3403546.

[42] Florian Heller and Jan Borchers. 2014. AudioTorch: using a smartphone as directional microphone in virtual audio spaces. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*. 483–488.

[43] Florian Heller and Jan Borchers. 2015. Audioscope: Smartphones as directional microphones in mobile audio augmented reality systems. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 949–952.

[44] Florian Heller and Johannes Schöning. 2018. Navigatone: Seamlessly embedding navigation cues in mobile music listening. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–7.

[45] Thomas Hermann, Anselm Venezian Nehls, Florian Eitel, Tarik Barri, and Marcus Gammel. 2012. Tweetscapes – Real-Time Sonification of Twitter Data Streams for Radio Broadcasting.

[46] Joo Young Hong, Jianjun He, Bhan Lam, Rishabh Gupta, and Woon-Seng Gan. 2017. Spatial Audio for Soundscape Design: Recording and Reproduction. *Applied Sciences* 7, 6 (2017). https://doi.org/10.3390/app7060627.

[47] Matthias Hoppe, Jakob Karolus, Felix Dietz, Paweł W. Woźniak, Albrecht Schmidt, and Tonja-Katrin Machulla. 2019. *VRsneaky: Increasing Presence in VR Through Gait-Aware Auditory Feedback*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3290605.3300776.

[48] Hartmut Ising and B. Kruppa. 2003. Health effects caused by noise: Evidence in the literature from the past 25 years. *Noise & health* 6 (11 2003), 5–13.

[49] Sampath Jayalath and Nimsiri Abhayasinghe. 2013. A gyroscopic data based pedometer algorithm. In *2013 8th International Conference on Computer Science Education*. 551–555. https://doi.org/10.1109/ICCSE.2013.6553971.

[50] Jon Kabat-Zinn. 1982. An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: Theoretical considerations and preliminary results. *General Hospital Psychiatry* 4, 1 (1982), 33–47. https://doi.org/10.1016/0163-8343(82)90026-3.

[51] Jon Kabat-Zinn. 2013. *Full catastrophe living, revised edition: how to cope with stress, pain and illness using mindfulness meditation*. Hachette uK.

[52] Jon Kabat-Zinn. 2017. Walking Meditations. *Mindfulness* 8, 1 (01 02 2017), 249–250. https://doi.org/10.1007/s12671-016-0638-1.

[53] Jian Kang, Francesco Aletta, Truls Gjestland, Lex Brown, Dick Botteldooren, Brigitte Schulte-Fortkamp, Peter Lercher, Irene van Kamp, Klaus Genuit, André Fiebig, José Coelho,

Luigi Maffei, and Lisa Lavia. 2016. Ten questions on the soundscapes of the built environment. *Building and Environment* 108 (08 2016). https://doi.org/10.1016/j.buildenv.2016.08.011.
[54] Xiaomin Kang, Baoqi Huang, and Guodong Qi. 2018. A Novel Walking Detection and Step Counting Algorithm Using Unconstrained Smartphones. *Sensors* 18, 1 (2018). https://doi.org/10.3390/s18010297.
[55] Fahim Kawsar, Chulhong Min, Akhil Mathur, Alessandro Montanari, Utku Günay Acer, and Marc Van den Broeck. 2018. esense: Open earable platform for human sensing. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*. 371–372.
[56] Angelika C. Kern and Wolfgang Ellermeier. 2020. Audio in VR: Effects of a Soundscape and Movement-Triggered Step Sounds on Presence. *Frontiers in Robotics and AI* 7 (2020), 20. https://doi.org/10.3389/frobt.2020.00020.
[57] Caitlin R. Kight and J. Swaddle. 2011. How and why environmental noise impacts animals: an integrative, mechanistic review. *Ecology letters* 14 10 (2011), 1052–1061.
[58] Hiroki Kobayashi, Ryoko Ueoka, and Michitaka Hirose. 2008. Wearable Forest-Feeling of Belonging to Nature. In *Proceedings of the 16th ACM International Conference on Multimedia (MM '08)*. Association for Computing Machinery, New York, NY, USA, 1133–1134. https://doi.org/10.1145/1459359.1459600.
[59] Andreas Komninos, Mark D. Dunlop, David Rowe, Allan Hewitt, and Steven Coull. 2015. Using degraded music quality to encourage a health improving walking pace: BeatClearWalker. In *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*. IEEE, 57–64.
[60] Gregory Kramer, Bruce Walker, Terri Bonebright, Perry Cook, John Flowers, Nadine Miner, John Neuhoff, Robin Bargar, Stephen Barrass, Jonathan Berger, Grigori Evreinov, W. Fitch, M. Grohn, S. Handel, Hans Kaper, H. Levkowitz, S. Lodha, Barbara Shinn-Cunningham, Mary Simoni, and Sever Tipei. 1999. The sonification report: Status of the field and research agenda. Report prepared for the National Science Foundation by members of the International Community for Auditory Display. (01 1999).
[61] Mark Lawton, Stuart Cunningham, and Ian Convery. 2020. Nature soundscapes: an audio augmented reality experience. In *Proceedings of the 15th International Conference on Audio Mostly*. 85–92.
[62] Judith Ley-Flores, Frédéric Bevilacqua, Nadia Bianchi-Berthouze, and Ana Taiadura-Jiménez. 2019. Altering body perception and emotion in physically inactive people through movement sonification. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. 1–7. https://doi.org/10.1109/ACII.2019.8925432.
[63] Nicholas Mariette. 2013. Human factors research in audio augmented reality. In *Human factors in augmented reality environments*. Springer, 11–32.
[64] Mark McGill, Stephen Brewster, David McGookin, and Graham Wilson. 2020. Acoustic transparency and the changing soundscape of auditory mixed reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–16.
[65] David McGookin and Stephen Brewster. 2011. PULSE: An Auditory Display to Provide a Social Vibe. In *Proceedings of Interacting with Sound Workshop: Exploring Context-Aware, Local and Social Audio Applications (IwS '11)*. Association for Computing Machinery, New York, NY, USA, 12–15. https://doi.org/10.1145/2019335.2019338.
[66] David K. McGookin, Stephen A. Brewster, and Pablo Priego. 2009. Audio Bubbles: Employing Non-speech Audio to Support Tourist Wayfinding. In *HAID*.
[67] Fritz Menzer, Anna Brooks, Pär Halje, Christof Faller, Martin Vetterli, and Olaf Blanke. 2010. Feeling in control of your footsteps: conscious gait monitoring and the auditory consequences of footsteps. *Cognitive neuroscience* 1, 3 (2010), 184–192.
[68] Brett R. C. Molesworth, Marion Burgess, and Daniel Kwon. 2013. The use of noise cancelling headphones to improve concurrent task performance in a noisy environment. *Applied Acoustics* 74, 1 (2013), 110–115. https://doi.org/10.1016/j.apacoust.2012.06.015.
[69] Thomas Münzel, Mette Sørensen, Frank Schmidt, Erwin Schmidt, Sebastian Steven, Swenja Kröller-Schön, and Andreas Daiber. 2018. The adverse effects of environmental noise exposure on oxidative stress and cardiovascular risk. *Antioxidants & redox signaling* 28, 9 (2018), 873–908.
[70] Benjamin Müller, Andreas Liebl, and Noemi Martin. 2020. Is wearing active-noise-cancelling headphones in open space offices beneficial for cognitive performance and employee satisfaction? Results of two laboratory studies.
[71] Ayumi Ohnishi, Isao Nishiyama, Tsutomu Terada, and Masahiko Tsukamoto. 2019. An Auditory Feedback System to Improve the Foot Pressure Balance for Runners. In *Proceedings of the 17th International Conference on Advances in Mobile Computing & Multimedia (MoMM2019)*. Association for Computing Machinery, New York, NY, USA, 94–101. https://doi.org/10.1145/3365921.3365941.
[72] Nuria Oliver and Fernando Flores-Mangas. 2006. MPTrain: a mobile, music and physiology-based personal trainer. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*. 21–28.
[73] World Health Organization. 2010. *Global Recommendations on Physical Activity for Health*. World Health Organization. https://www.who.int/publications/i/item/9789241599979.
[74] World Health Organization et al. 2011. *Burden of disease from environmental noise: Quantification of healthy life years lost in Europe*. World Health Organization. Regional Office for Europe.
[75] World Health Organization et al. 2018. Environmental noise guidelines for the European region. (2018).
[76] Joseph Plazak and Marta Kersten-Oertel. 2018. A Survey on the Affordances of "Hearables". *Inventions* 3, 3 (2018), 48.
[77] Susaree Prakhinkit, Siriluck Suppapitiporn, Hirofumi Tanaka, and Daroonwan Suksom. 2014. Effects of Buddhism Walking Meditation on Depression, Functional Fitness, and Endothelium-Dependent Vasodilation in Depressed Elderly. *The Journal of Alternative and Complementary Medicine* 20, 5 (2014), 411–416. https://doi.org/10.1089/acm.2013.0205 arXiv:https://doi.org/10.1089/acm.2013.0205 PMID: 24372522.
[78] Josephine Reid, Erik Geelhoed, Richard Hull, Kirsten Cater, and Ben Clayton. 2005. *Parallel Worlds: Immersion in*

*Location-Based Experiences*. Association for Computing Machinery, New York, NY, USA, 1733–1736. https://doi.org/10.1145/1056808.1057009.

[79] Matthew W. M. Rodger, William R. Young, and Cathy M. Craig. 2014. Synthesis of Walking Sounds for Alleviating Gait Disturbances in Parkinson's Disease. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 22 (2014), 543–548.

[80] Chiara Rossitto, Louise Barkhuus, and Arvid Engström. 2016. Interweaving place and story in a location-based audio drama. *Personal and Ubiquitous Computing* 20, 2 (2016), 245–260.

[81] Nitin Sawhney and Chris Schmandt. 2000. Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM transactions on Computer-Human interaction (TOCHI)* 7, 3 (2000), 353–383.

[82] R. Murray Schafer. 1993. *The soundscape: Our sonic environment and the tuning of the world*. Simon and Schuster.

[83] Mariana Seiça, Buga Lopes, Pedro Martins, and Amílcar Cardoso. 2017. Sonifying Twitter's Emotions Through Music.

[84] S. Serafin, L. Turchet, and R. Nordahl. 2011. Auditory feedback in a multimodal balancing task: Walking on a virtual plank. *Proceedings of the 8th Sound and Music Computing Conference, SMC 2011* (2011).

[85] Mark Shepard. 2006. Tactical sound garden [tsg] toolkit. In *3rd international workshop on mobile music technology*.

[86] Mark Shepard. 2007. Tactical Sound Garden Toolkit. In *ACM SIGGRAPH 2007 Art Gallery (SIGGRAPH '07)*. Association for Computing Machinery, New York, NY, USA, 219. https://doi.org/10.1145/1280120.1280176.

[87] Barbara G. Shinn-Cunningham and Virginia Best. 2008. Selective Attention in Normal and Impaired Hearing. *Trends in Amplification* 12, 4 (2008), 283–299. https://doi.org/10.1177/1084713808325306 arXiv:https://doi.org/10.1177/1084713808325306 PMID: 18974202.

[88] Marjan Sikora, Mladen Russo, Jurica Undefinederek, and Ante Jurčević. 2018. Soundscape of an Archaeological Site Recreated with Audio Augmented Reality. *ACM Trans. Multimedia Comput. Commun. Appl.* 14, 3, Article 74 (July 2018), 22 pages. https://doi.org/10.1145/3230652.

[89] Hyungki Son, Hyunjae Gil, Sangkyu Byeon, Sang-Youn Kim, and Jin Ryong Kim. 2018. Realwalk: Feeling ground surfaces while walking in virtual reality. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–4.

[90] Michael Southworth. 1969. The Sonic Environment of Cities. *Environment and Behavior* 1, 1 (1969), 49–70. https://doi.org/10.1177/001391656900100104 arXiv:https://doi.org/10.1177/001391656900100104.

[91] Nicole Rachel Srednicki. 2017. *Exploring a Change in Meditation Practice Using Noise-Canceling Headphones*. Georgetown University.

[92] Ana Tajadura-Jiménez, Maria Basia, Ophelia Deroy, Merle Fairhurst, Nicolai Marquardt, and Nadia Bianchi-Berthouze. 2015. As Light as Your Footsteps: Altering Walking Sounds to Change Perceived Body Weight, Emotional State and Gait. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 2943–2952. https://doi.org/10.1145/2702123.2702374.

[93] Yuichiro Takeuchi. 2010. Gilded gait: reshaping the urban experience with augmented footsteps. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. 185–188.

[94] Robert Triggs. 2021. Android's Bluetooth latency needs a serious overhaul. https://www.soundguys.com/android-bluetooth-latency-22732/ Online; accessed 12 June 2022.

[95] Luca Turchet, Ivan Camponogara, and Paola Cesari. 2015. Interactive footstep sounds modulate the perceptual-motor aftereffect of treadmill walking. *Experimental brain research* 233, 1 (2015), 205–214.

[96] Luca Turchet and Stefania Serafin. 2013. Investigating the amplitude of interactive footstep sounds and soundscape reproduction. *Applied Acoustics* 74, 4 (2013), 566–574. https://doi.org/10.1016/j.apacoust.2012.10.010.

[97] Luca Turchet, Stefania Serafin, and Paola Cesari. 2013. Walking Pace Affected by Interactive Sounds Simulating Stepping on Different Terrains. *ACM Trans. Appl. Percept.* 10, 4, Article 23 (2013), 14 pages. https://doi.org/10.1145/2536764.2536770.

[98] Loes van Renswouw, Jelle Neerhof, Steven Vos, Pieter van Wesemael, and Carine Lallemand. 2021. *Sensation: Sonifying the Urban Running Experience*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411763.3451788.

[99] Timothy Van Renterghem, Kris Vanhecke, Karlo Filipan, Kang Sun, Toon De Pessemier, Bert De Coensel, Wout Joseph, and Dick Botteldooren. 2020. Interactive soundscape augmentation by natural sounds in a noise polluted urban park. *Landscape and Urban Planning* 194 (2020), 103705. https://doi.org/10.1016/j.landurbplan.2019.103705.

[100] Yolanda Vazquez-Alvarez, Ian Oakley, and Stephen A. Brewster. 2012. Auditory Display Design for Exploration in Mobile Audio-Augmented Reality. *Personal Ubiquitous Comput.* 16, 8 (Dec. 2012), 987–999. https://doi.org/10.1007/s00779-011-0459-0.

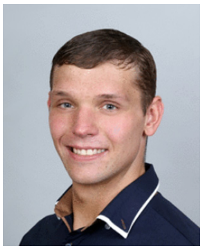
[101] Julius von Willich, Martin Schmitz, Florian Müller, Daniel Schmitt, and Max Mühlhäuser. 2020. *Podoportation: Foot-Based Locomotion in Virtual Reality*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376626.

[102] Bob G. Witmer and Michael J. Singer. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments* 7, 3 (06 1998), 225–240. https://doi.org/10.1162/105474698565686 arXiv:https://direct.mit.edu/pvar/article-pdf/7/3/225/1836425/105474698565686.pdf.

[103] Hong Ying, C. Silex, A. Schnitzer, S. Leonhardt, and M. Schiek. 2007. Automatic Step Detection in the Accelerometer Signal. In *4th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2007)*, Steffen Leonhardt, Thomas Falck, and Petri Mähönen (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 80–85.

[104] Andreas Zimmermann and Andreas Lorenz. 2008. LISTEN: A User-Adaptive Audio-Augmented Museum Guide. *User Modeling and User-Adapted Interaction* 18, 5 (2008), 389–416. https://doi.org/10.1007/s11257-008-9049-x.

## Bionotes



**Maximilian Schrapel**
Leibniz University Hannover, HCI Group, Hannover, Germany
**maximilian.schrapel@hci.uni-hannover.de**

Maximilian Schrapel is a Ph. D. student at the Leibniz University Hannover, Germany. He received a Bachelor's degree in electrical engineering from the Hochschule Hannover and a Master's degree in computer science from the Leibniz University Hannover, Germany. He currently works at the Human-Computer Interaction Group led by Michael Rohs. Maximilian's research interests include surface computing, mobile systems, augmented reality and sensor technology. For his dissertation, he investigates how ordinary, ubiquitous surfaces can be integrated into human-computer interactions.

**Janko Happe**
Leibniz University Hannover, HCI Group, Hannover, Germany
**janko.happe@gmail.com**

Janko Happe is a software engineer and mobile application developer. He received a Master's degree in computer science from the Leibniz University Hannover, Germany.

**Michael Rohs**
Leibniz University Hannover, HCI Group, Hannover, Germany
**michael.rohs@hci.uni-hannover.de**

Michael Rohs is a professor of human-computer interaction in the Faculty of Electrical Engineering and Computer Science at Leibniz Universität Hannover, Germany. His primary research interests are in mobile human-computer interaction and pervasive computing. His work focuses on novel interaction techniques for mobile devices, the use of sensors for mobile interactions, mobile haptic feedback, and the integration of physical and virtual resources in human-computer interactions. He received a Diplom in computer science from the Technische Universität Darmstadt, Germany, a Master's degree in computer science from the University of Colorado at Boulder, USA, and a Ph. D. in computer science from ETH Zurich, Switzerland. In the past, he worked as a research assistant in the Distributed Systems Group at ETH Zurich, as a senior research scientist at Deutsche Telekom Laboratories and Technische Universität Berlin, Germany, and as an Assistant Professor at the University of Munich.