Manuscript accepted for publication in Personal and Ubiquitous Computing.
The final publication is available at Springer via http://dx.doi.org/10.1007/s00779-016-0906-z

Guided by music: pedestrian and cyclist navigation with route and beacon guidance

Robert Albrecht · Riitta Väänänen · Tapio Lokki

Abstract Music listening and navigation are both common tasks for mobile device users. In this study, we integrated music listening with a navigation service, allowing users to follow the perceived direction of the music to reach their destination. This navigation interface provided users with two different guidance methods, route guidance and beacon guidance. The user experience of the navigation service was evaluated with pedestrians in a city centre and with cyclists in a suburban area. The results show that spatialized music can be used to guide pedestrians and cyclists towards a destination without any prior training, offering a pleasant navigation experience. Both route and beacon guidance were deemed good alternatives, but the preference between them varied from person to person and depended on the situation. Beacon guidance was generally considered to be suitable for familiar surroundings, while route guidance was seen as a better alternative for areas that are unfamiliar or more difficult to navigate.

Keywords Navigation · Walking · Cycling · Music · Spatial audio

1 Introduction

Many people listen to music with headphones when on the move, e.g., when walking or cycling. People often travel the same routes, but sometimes they need to find their way to a new destination. Looking at a map or navigation instructions on a screen takes attention away from the environment, and is cumbersome to do while walking and especially while cycling. Thus, a hands- and eyes-free user interface for navigation is a safer and more convenient alternative. For this task, an auditory user interface could be utilized, presenting verbal turn-by-turn instructions or other sounds to guide the user. Such additional sounds might, however, disturb listening to the music, and be unclear unless the music is attenuated when they are presented. Instead, directional cues could be added to the music itself, allowing the user to follow the direction of the music to reach the destination [1–5].

In addition to the user interface employed, be it visual, auditory, tactile, or multimodal, navigation systems differ in what type of instructions they give to the user. Turn-by-turn instructions are popular and effective for automobile navigation, since car drivers are limited to driving along a road network. Pedestrians, on the other hand, are often not as restricted when it comes to choosing their route. They can move more or less freely over parks and squares, and do not need to take one-way traffic into account. Additionally, pedestrians can stop at any moment to decide which way to continue, something drivers often cannot do.

Based on these differences, beacon guidance has often been suggested as a viable method for pedestrian navigation systems [1–3, 6–8]. With turn-by-turn guidance, the navigation system performs the route-finding, whereas with beacon guidance, the system conveys the direction straight to the destination, and the navigation task is left to the user. Alternatively, several beacons, or waypoints, can be used in succession in cases where more detailed guidance around obstacles is needed. Beacon guidance can be implemented, e.g., using spatial audio or tactile interfaces that inform the user about the direction of the destination. One advantage of beacon guidance, compared with turn-by-turn guidance, is the simplicity of implementation: no information about road

The research leading to these results has received funding from Nokia Technologies.

R. Albrecht · T. Lokki
Aalto University, Department of Computer Science
Espoo, Finland
E-mail: robert.albrecht@aalto.fi

R. Väänänen
Nokia Technologies
Espoo, Finland

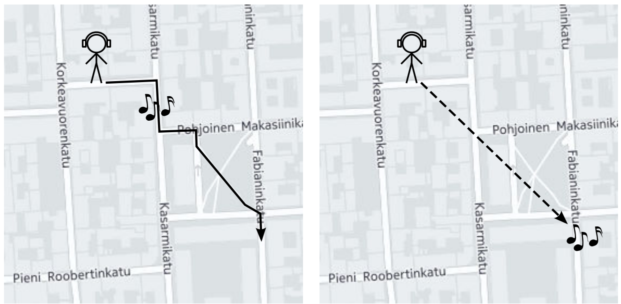
networks is needed, only the coordinates of the destination and the user.

A third alternative is what we call route guidance. This type of navigation guidance has previously been proposed (but not investigated) by Strachan et al. [2], who suggested leading the user along a route with the audio source a certain distance ahead like a "carrot on a stick." With this method the user can be directed, e.g., along an arbitrary route across a large open field. In this type of situation, it would be difficult to provide the user with exact verbal instructions. Compared with beacon guidance with multiple waypoints, route guidance is a more seamless method, where the audio source constantly moves along the route some distance ahead of the user instead of "jumping" from waypoint to waypoint. Only little prior research exists on route guidance: Väänänen et al. [9] evaluated route guidance with the user following the sound of a horse walking on the route ahead. Instead of using spatial audio, route guidance can also be implemented using tactile interfaces.

Utilizing a prototype navigation service, this study investigates the use of both route and beacon guidance for pedestrian and cyclist navigation, where the user is guided by music heard from the direction where he or she should be going, as illustrated in Fig. 1. The following scenario illustrates use of such a navigation service:

> Jessica leaves work, puts her headphones on, and listens to music as she usually does. She is going to meet a friend at a restaurant across town. As she is unfamiliar with that part of town, she turns on a navigation application, chooses beacon guidance mode, and selects the destination. Now she can hear the music coming from a direction behind her, so she turns around and starts walking in that direction. She finds her own route through town, but the music constantly informs her of the direction that she should be walking in. Halfway across town, Jessica looks at the clock and realizes that she already is in a bit of a hurry. As she feels slightly unsure about the fastest route to the restaurant, she switches to route guidance mode, and now she can just follow the direction of the music at every intersection to get to the restaurant in time to meet her friend.

Previous studies on music guidance for navigation have utilized stereo balance adjustment to convey the direction to the destination or the next waypoint [1–4]. For example, if you hear the music in the right ear, you should be heading right, and if you hear the music equally loud in both ears, you should head straight forward, or possibly backward. In our navigation service, we instead decided to present the music using head-related transfer functions (HRTFs), which provide a more natural representation of directions



(a) Route guidance (b) Beacon guidance

Fig. 1 Illustrations of the two guidance types tested in this study

and thereby make it easier, e.g., to distinguish between directions forward and backward.

To be able to guide the user in the right direction, a navigation system must know which direction the user is facing. Previous studies on music guidance have used the compass of a hand-held device [1, 2] or the direction of motion of the user [3] as a reference for presenting the direction where the user should go. However, to present the music in the correct direction irrespective of how the user turns his or her head, it is the orientation of the head that is of interest. Smartphones contain the necessary orientation sensors, but the phone must be attached to the head of the user in order to utilize them for head tracking [4]. We instead decided to utilize a separate head tracker with our navigation service.

In this study, we extend the work on route guidance by Väänänen et al. [9] (presented in Sect. 2.3.2), by using music as the guiding sound, comparing route guidance with beacon guidance, and investigating both pedestrian and cyclist navigation. Compared with previous studies on music guidance for navigation [1–4] (presented in Sect. 2.3.3 and 2.4), the current study differs in the inclusion of route guidance, utilization of HRTFs for spatial audio, and head orientation tracking (used only in [4]). This study also aims to provide a more extensive look into many of the details, e.g., safety, ease of use, and cognitive load, related to the user experience of music guidance in a real environment.

The rest of this paper continues with a literature review on spatial audio guidance for navigation as well as safety considerations. We then describe our prototype navigation system and the three user studies conducted with it. The first user study was done with pedestrians in a city centre. In the second study, a further developed version of the navigation system was again tested by pedestrians. In the third study, cyclists evaluated the navigation system in a suburban area. We continue by presenting the results of the user studies, followed by discussion and design recommendations based on these findings.

2 Background and related research

This section provides background on sound localization and the use of head orientation tracking, and presents previous laboratory experiments and field studies on spatial audio guidance for pedestrian and cyclist navigation. Unless otherwise stated, these studies were performed with headphones. Studies on safety issues related to the use of headphones and mobile phones when walking or cycling are also discussed.

2.1 Sound localization

Spatial audio guidance methods for navigation take advantage of the natural ability of humans to localize the source of a sound. In headphone reproduction, spatial sound localization cues can be simulated using digital signal processing. The main cues utilized for horizontal localization are the interaural time difference (ITD) and interaural level difference (ILD) [10]. These differences are caused by the shadowing effect of the head and the difference in travel time to the two ears from sound sources that are not located in the median plane. The simplest way to alter the perceived azimuth of a virtual sound source is to apply an ILD by adjusting the relative amplitude of the sound presented to each ear. This method is referred to as amplitude panning, and stereo balance adjustment is an example of it. Alternatively, or additionally, a time difference (ITD) may be introduced between the sounds presented to the two ears.

In addition to the simple ILD and ITD, humans utilize other cues for sound localization, especially for vertical localization. The pinna, in particular, affects the temporal and spectral characteristics of the sound entering the ear canal through reflections, shadowing, dispersion, diffraction, interference and resonance [10]. A head-related transfer function (HRTF) describes these effects, in addition to the ILD and ITD, related to a specific sound source position. HRTFs modeled or measured using real or dummy heads can thus be applied to sound signals presented over headphones to make the sound appear as coming from the desired direction.

ILDs and ITDs alone do not provide the means to distinguish between a sound coming from the front and a sound coming from the back, since there is no difference in the ILDs and ITDs of, e.g., a sound source straight ahead and a sound source straight behind. In fact, a sound presented using ILDs and ITDs is likely to be perceived as originating inside the head on a line between the ears. HRTFs provide cues which, to some extent, help distinguish between sounds coming from the front and sounds coming from the back, and also aid in making the sounds appear to originate from outside the head [11].

2.2 Head orientation tracking

Using spatial audio techniques with headphones, virtual sound sources can be presented so that they are heard from different directions. To be able to place these sound sources so that they are stable with respect to the real world around the user, the orientation of the user, and specifically the head of the user, must be known. Otherwise the sound sources will rotate, with respect to the real world, as the user rotates his or her head.

In experimental spatial audio systems, head orientation tracking (or simply head tracking) is often implemented using a separate head tracker unit. Although not currently available, headphones with integrated orientation tracking (e.g. the Jabra Intelligent Headset[1]) can soon be expected on the consumer market. Orientation sensors are also typically included in smartphones, which means that a smartphone can be attached to the headphones and used for orientation tracking [4], but this approach will probably not become popular.

Front-back confusion is a common problem with spatial audio using headphones. That is, sound sources that should be heard from the front are heard from behind, and vice versa. As mentioned earlier, the use of HRTFs, compared with simpler spatial audio techniques, will provide better front-back discrimination. A further improvement can be achieved by the use of head tracking. With low-latency head tracking, sound sources remain stable as the user rotates his or her head, which helps to resolve front-back confusion and thereby improve sound source localization [12].

If head tracking is unavailable, users cannot rely on head movements to help localize sound sources. Instead, users may develop other techniques to improve localization. For example, if the orientation of a mobile device is tracked, the user can rotate this device to pinpoint the direction of the sound source [1], as described in Sect. 2.3.3. Even though such techniques might prove effective, they do not provide the natural listening experience that is possible with head tracking.

2.3 Spatial audio guidance

The use of spatial audio for navigational guidance has previously been investigated under different conditions. Here, we present the results from relevant laboratory experiments, as well as field studies on using spatial audio guidance or music guidance.

2.3.1 Laboratory experiments on spatial audio guidance

Klatzky et al. [13] compared spatial language with spatial sound for navigation tasks without vision. In the spatial lan-

[1] http://intelligentheadset.com

guage mode, participants were given the verbal instructions "left" and "right" to indicate the correct heading towards the next waypoint, and "straight" if the waypoint was in front of them. In the spatial sound mode, head tracking was used, and the direction of a beeping sound, presented using HTRFs, indicated the correct heading. In the absence of additional cognitive load, there was no performance difference between the two modes, but when participants had to perform an additional cognitively demanding task, they performed better in the spatial sound mode with respect to distance travelled and time spent. In the spatial sound mode, the additional task performance was also better. The results suggest that spatial sound is a superior alternative for navigation tasks, producing relatively low cognitive load and leaving more resources for other tasks, such as observing the environment.

A few studies have compared different spatial audio techniques for navigation in virtual environments. The results of a study by Lokki et al. [14] indicate that using ITDs and ILDs together gives sufficient cues for navigation, and this method actually resulted in significantly faster completion times than a method using HRTFs. Using ITDs alone resulted in inferior error rates, but completion times comparable to the ones with HRTFs.

Larsen et al. [15] compared the use of ILDs with the use of HRTFs. The HRTF method proved significantly better in terms of localization precision and speed, as well as navigation performance.

While Lokki et al. and Larsen et al. studied navigation in an open virtual space, Gonot et al. [16] used a virtual environment consisting of a complex road network. Participants were asked to locate, in a given order, nine sound sources simultaneously emitting different sounds. The results showed a better performance, measured by the decision time needed at crossroads, with HRTFs than with a stereo technique providing ITD and ILD cues. Beacon guidance with waypoints resulted in better performance than guidance with a single beacon at the destination. These results, correlated, to some extent, with subjective workload ratings.

Tran et al. [17] investigated the use of different sounds for spatial audio navigation tasks, both with loudspeakers and with headphones using HRTFs. The results showed high negative correlation between localization error and subjective quality ratings. Continuous sounds were preferred over pulsed; participants complained about "losing track" of the location of pulsed sounds. The results also suggest that sounds used for navigation should be relatively wideband. Additionally, Tran et al. suggest that the sounds should be "artificial," so that they do not blend into the environment. These results indicate that music should be well-suited for spatial audio navigation tasks.

Fujimoto and Turk [5] studied the use of different combinations of continuous audio, specifically music, and tactile cues for navigation. The study was limited to interfaces that are possible with off-the-shelf smartphones and usable with the phone in the pocket. Direction was communicated either through vibration patterns or spatial audio (the exact technique not defined) with either a stationary or moving sound source. Distance was represented either by the time between vibrations or by the audio volume. No orientation tracking was used, and only directions in front of the user were presented.

A total of 41 people participated in the study, which took place in a laboratory environment. Based on the results, conveying direction by stationary spatial audio seems a better alternative than using a sweeping motion to emphasize the direction, as users preferred the stationary alternative and found it, e.g., less annoying. In addition, the error in distance estimation was greater if the direction was conveyed by sweeps. Spatial audio and vibrations, however, were deemed equally good alternatives for conveying direction. Audio volume was considered a better representation of distance than the time between vibrations, being less annoying and frustrating, easier to use, as well as less mentally demanding.

Of the interfaces tested, two combinations proved the best alternatives. One was using spatial audio for direction and volume for distance, the other was using vibrations for direction and volume for distance. However, a few participants commented that the multimodal interfaces, such as the one using vibrations and volume, added some cognitive load.

2.3.2 Pedestrian navigation using spatial audio guidance

Numerous navigation systems using spatial audio guidance have previously been suggested and evaluated. Here, we summarize studies by Calvo et al. comparing auditory, tactile, and visual displays for beacon guidance with waypoints. We also present prior work on route guidance by Väänänen et al., which the current study is a further development of.

In two different studies, Calvo et al. evaluated auditory and tactile navigation displays, comparing them with map-based navigation. In the first user study [6], an HRTF-based display and a vibration-belt display were tested, both with a 45-degree resolution and displaying the direction of the following waypoint relative to the compass heading of a handheld mobile phone. These displays were compared with an allocentric map showing the position of the participant and the next waypoint. Twelve participants tested all three display types on one-kilometer long routes on a university campus. The results showed no difference in completion times or navigation errors between the display types.

In the second user study [7], the auditory and tactile displays were compared with both an allocentric and an egocentric map, and with a visual arrow with the same

45-degree resolution. In this study, ten participants tested the navigation display types on 800-meter long routes in an open field. The completion times were significantly shorter using the egocentric map than with the allocentric map or tactile display. The auditory display provided slightly faster navigation than the tactile display. The study suggested that both auditory and tactile displays can be used effectively for navigation without visual information from either a display or the environment, with high usability and low mental workload.

Väänänen et al. [9] evaluated route guidance with a previous version of the navigation system investigated in this study. A virtual audio guide, represented by the sound of a walking horse, lead the way on the route a certain distance ahead of the user. The sound was presented using head tracking and HRTF processing. Additionally, a car turning signal was used to emphasize a forthcoming turn. Acoustically transparent headphones, utilizing a microphone hear-through technique [18], were used to allow users to hear the environment naturally. In this user study, ten participants tested the system on a 500-meter training route and a 900-meter test route in a city environment.

All but one of the participants were able to follow the audio guidance to the destination. Most participants felt that they were walking safely because they could hear the traffic and keep their phone in the pocket, and were thereby able to follow the guidance while paying attention to the surroundings at the same time. During the tests, the experimenters initiated a discussion with the participants, and additionally participants listened to music during a part of the route. It varied greatly how easy and pleasant participants found it to have a discussion at the same time as listening to the guidance, traffic, and possibly music. Based on the results, Väänänen et al. suggested that users should be able to choose from different sound themes, and whether the guidance is audible all the time or not.

2.3.3 Pedestrian navigation using spatial music guidance

While many experimental navigation systems with audio guidance use specific sounds explicitly for navigation tasks, adapting audio that the user otherwise also would listen to has been suggested. Especially music is something that many pedestrians listen to while walking around in a city, for example.

The ONTRACK beacon guidance system by Jones et al. [1] adjusted the stereo balance of music to guide users to the next waypoint. The panning was performed with a 30-degree resolution, based on the compass heading relative to the direction of the waypoint. In an initial lab prototype of ONTRACK, the volume of the music was adjusted so that it increased as users approached the next waypoint, and decreased if they were moving away from it. In a user study, this distance cue was found to be much too subtle, and users were confused as to whether they were walking towards or away from the waypoint. As Jones et al. also note, one problem with simply adjusting the left-right balance of music to indicate a direction is that some music tracks use stereo panning for effect. If there is prominent panning to either direction in the track, this might be mistaken for navigation cues.

A few participants in the field trial of ONTRACK developed an effective technique for "probing" the direction to the next waypoint. By moving the GPS device containing the compass from side to side and listening to the correspondingly changing panning cues, the participants were able to confirm that they were moving in the right direction. Overall, the results of the lab and field user studies indicated that the concept could lead to both useful and enjoyable applications.

A similar system for beacon guidance with music, gpsTunes, was developed by Strachan et al. [2]. The direction to the next waypoint was presented by amplitude panning of the music, and the distance was conveyed by changing the volume. Like ONTRACK, gpsTunes did not use head tracking, but instead relied on hand-held sensors for heading and location. The system was briefly evaluated with four users on a sports field, showing it to be a promising concept.

Melodious Walkabout by Etter and Specht [3] is another implementation of beacon guidance using music or other audio content. Stereo balance adjustment was here used to present the direction to the destination, relative to the user's current direction of motion. To indicate that the destination was behind the user, the music was low-pass filtered. When the user was getting close to the destination, the volume of the music gradually decreased.

Melodious Walkabout was evaluated with 24 participants on two different routes, with the order of the routes counterbalanced. Since the participants were not told how the guidance worked, a strong learning effect was observed, and both the time taken and the distance covered to complete the second route was significantly shorter than for the first route. For one of the routes, participants listened to a favourite song that they were asked to bring, while a song selected by the experimenters was used for the other route. The music played did not have a significant impact on the participants' performance. However, listening to their own music had a significant and positive impact on how clearly the participants thought that they heard the navigation cues and how satisfied they were with the navigation. Overall, participants thought that the navigation worked well and required minimal attention.

2.4 Cyclist navigation

Compared with pedestrian navigation, the need for eyes-free and hands-free navigation systems is much more prominent when it comes to cyclist navigation. Both hands are typically occupied during cycling and the higher speeds during cycling require more attention to the surroundings than is needed when walking. However, little research has been done concerning cyclist navigation.

Zwinderman et al. [4] proposed and briefly evaluated a system using beacon guidance for cyclist navigation. In the implemented prototype, the stereo balance of music was modified to indicate the direction of the destination. The orientation of the user's head was provided by the magnetometer of a smartphone that was attached to the headphones worn by the user.

Pielot et al. [19] suggested that when it comes to touristic exploration, simply conveying the general direction of the destination is a viable approach, but pointed out that this direction has to be constantly available, in contrast to turn-by-turn instructions. The challenge is to avoid distracting the user, by, e.g., requiring him or her to periodically check the display of a smartphone. Pielot et al. proposed using tactile stimuli rather than audio for this purpose, since listening to the environment often is an important part of a touristic experience, and audio might also become inaudible in noisy environments.

The Tacticycle prototype navigation system [19] was designed for exploratory bicycle trips. A smartphone was attached to the handlebar of the bicycle, showing a map where the user could see different points of interest and choose one of them as the destination. The system then provided beacon guidance towards the destination, by means of vibration motors attached to the handlebars. The relative intensity of the vibration of the left and right handlebars was used to convey the direction. Users were also notified about nearby points of interest through the tactile interface.

In a user study, participants found the tactile cues provided by Tacticycle to be unobtrusive. Participants felt oriented using the system, as well as encouraged to explore the environment. They also liked the eyes-free experience. Some of the participants, however, struggled with the idea of not having turn-by-turn instructions, since this is what usually is expected from a navigation system. Other participants, on the other hand, liked the freedom that it gave them.

2.5 Safety

When it comes to navigation in areas with traffic, safety is naturally a primary concern. There are two important aspects related to the safety of music guidance. One is related to the safety of music guidance compared with traditional navigation systems. The other is related to safety concerns when using headphones and listening to music while walking or cycling in traffic.

Lichenstein et al. [20] explain the danger of using headphones as a pedestrian as a result of two phenomena: inattentional blindness and environmental isolation. Inattentional blindness, which is due to the cognitive distraction of interpreting auditory stimuli and possibly manipulating electronic devices, reduces mental resource allocation or attention to outside stimuli. Environmental isolation is the inability to hear sounds from the surroundings due to attenuation caused by the headphones and masking caused by the sounds reproduced with the headphones.

Using audio output instead of visual output in a navigation system allows visual attention to remain on the traffic and surroundings, and should thereby increase safety, a notion supported by a study on driver behaviour and performance by Jensen et al. [21]. 30 participants in a between-subjects experiment tested audio, visual, and audio-visual output of a state-of-the-art car navigation system in real-world driving tasks. Visual output resulted in significantly worse driving performance than audio output, with more speeding violations and lateral control errors. Surprisingly, there were no significant differences in driving performance between visual and audio-visual output, despite the fact that audio-visual output significantly reduced eye glances at the navigation device.

Studies on pedestrian safety related to mobile phone use and music listening have provided somewhat contradictory results. Thompson et al. [22] observed 1102 pedestrians at 20 intersections in the USA. Text messaging, talking on the phone or talking with a companion increased the time it took to cross the intersection, while listening to music was associated with a decreased crossing time. Pedestrians writing text messages were much more likely to display unsafe crossing behaviour than undistracted pedestrians. Texting, listening to music, and talking in a group were all associated with an increased likelihood not to look both ways before crossing the street, while talking on the phone was not.

Walker et al. [23], on the other hand, found that male pedestrians listening to music more often looked left and right than those not listening to music at two observed crosswalks in Canada. No difference was, however, observed between female pedestrians listening or not listening to music. In this study, a total of 264 pedestrians were observed. Listening to music did not affect if people stopped or slowed down before crossing.

Nasar et al. [24] also made observations contradictory to those by Thompson et al. A total of 127 pedestrians were observed at three crosswalks in the USA, with those talking on the phone displaying a higher percentage of unsafe behaviour than those listening to music or not using any technology. The music listeners did not display significantly more unsafe behaviour than pedestrians not using any technology.

De Waard et al. [25] studied mobile phone use in the Netherlands while cycling, by means of observation, a questionnaire to accident-involved cyclists, and a cycling experiment. In the experiment, having a phone call resulted in reduced speed, reduced peripheral vision, and increased mental effort and risk ratings. Text messaging had a larger negative impact on the cycling performance, with an increased need for lateral space. Listening to music while cycling did not affect performance, but participants themselves associated it with a slightly increased risk. The observational and questionnaire studies did not, however, show any increased risk of accidents associated with phone use while cycling.

Although the mentioned studies present somewhat differing results, some conclusions can be drawn about the implications for pedestrian and cyclist navigation. Operating a mobile phone, e.g., texting or looking at a map or navigation instructions, while walking or cycling results in cognitive distraction, which presumably increases the risk of accidents. Using audio output instead of visual output allows visual attention to be focused on the environment and thereby improves safety. Listening to music does not seem to be a source of much cognitive distraction, but it does reduce the ability to hear sound from the environment that might warn about imminent dangers.

3 Navigation system prototype

The test platform we used in this study was developed for mobile pedestrian navigation in a city environment using spatial audio guidance. The first functional version of this platform was used in the earlier navigation user experience tests reported by Väänänen et al. [9] (summarized in Sect. 2.3.2). Based on the results of those tests, the platform was further developed and tuned to support navigation with the help of spatially processed music, and cyclist navigation in addition to pedestrian navigation.

3.1 Hardware platform

The navigation hardware setup used in the user studies is illustrated in Fig. 2. The development of the hardware setup was driven by the requirement of sufficient computing power to enable precise and low-latency head-tracked spatial audio, and updating the map and user position data in real time. We also aimed at building a setup, which would be as lightweight and unnoticeable to wear as possible. To reach these goals we used a tablet PC with the Windows 8 operating system, in-ear headphones, and an InterSense Wireless InertiaCube3 head orientation tracking device. The InertiaCube3 uses magnetometer, gyroscope and accelerometer sensors to track the orientation at a high resolution and accuracy, and communicates with the computing unit with sufficiently low latency. The total latency between the user's head movements and the corresponding change in the audio stimulus was approximately 100 ms, which should be small enough not to noticeably affect sound localization [26].

For tracking the position of the user, we used a Windows Phone device (either Nokia Lumia 620 or 1520). The phone was also used for controlling the navigation task and for viewing the progress of the user with respect to the map. For this reason, the phone was carried by the experimenter, who tried to stay close to the test participant during the tests, so that the GPS position represented the position of the participant as well as possible with this setup.

For reproducing the guidance-related sounds to the user in our pedestrian navigation studies, we used Nokia WH-208 in-ear headphones. We glued the earpieces of the headphones on a band that went over the user's head, and attached the head tracker on top of the headband. In our cyclist navigation study, we could not use a headband, but instead used a Nokia Purity WH-920 in-ear headset and attached the head tracker on top of the bicycle helmet that the participants wore.

In order to make the headphones acoustically transparent, we connected them to an audio augmented reality mixer, following the analog signal processing design principles of Tikander et al. [18]. The hear-through property of the headphones was obtained by placing an omnidirectional microphone outside each earpiece and feeding the microphone signal to the earpiece on the same side through the equalizer of the mixer. This allowed users to hear sounds from the surroundings, such as approaching cars, clearly while wearing the headphones.

The mixer included a USB audio interface for connecting it to the tablet PC. Through the USB interface, it was also possible to adjust the gain of the microphone signals before they were fed to the headphones, thus affecting how much the user heard of the environment. With this kind of system, users may block out background noise, or choose to listen to the environment at a normal or even amplified level, depending on the situation.

3.2 Software

The software in the prototype system was developed using C/C++ in a Windows 8 environment and utilizes the HERE map service to calculate routes and show maps. The main idea in our spatial audio guidance system is that the audio guide is represented by an HRTF-processed 3D audio object that conveys the direction where the user should go. As explained in Sect. 1, there are two approaches to placing

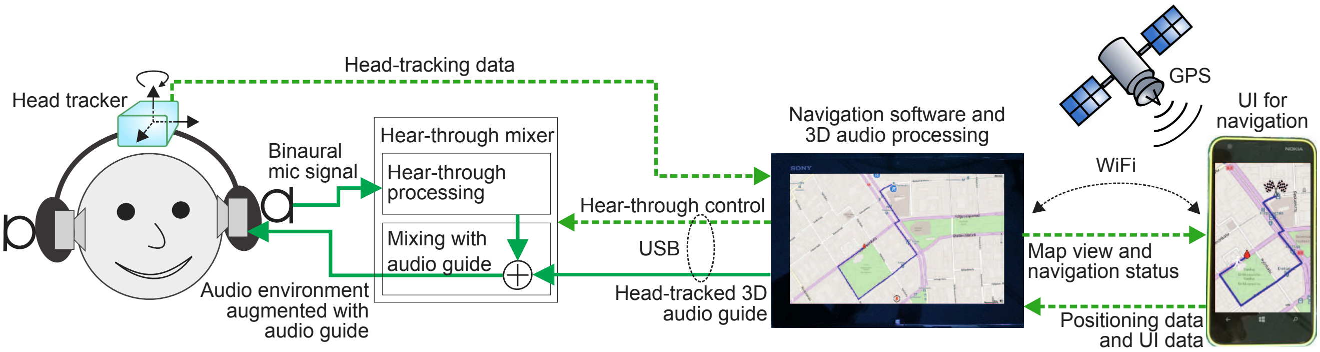
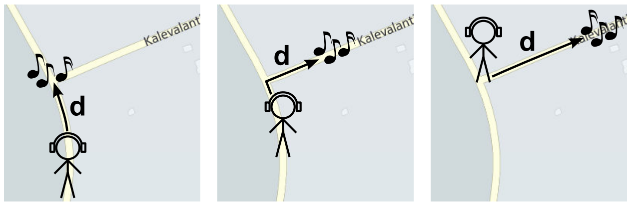


Fig. 2 The hardware setup used in the user studies

the virtual audio guide in the world surrounding the user. In route guidance mode, the guide is placed on the route ahead of the user so that it leads the way to the navigation destination, as shown in Fig. 1(a). In other words, the user finds his or her way by following the virtual guide. In beacon guidance mode, the 3D audio object is placed at the final destination, so that it acts as a beacon or a landmark, whose direction the user is constantly aware of, as illustrated in Fig. 1(b).

Although the distance between the user and the audio guide varies, this distance is not made audible to the user. However, a small and constant amount of artificial reverberation is added to help make the sound appear to originate outside the head of the user [12]. For a discussion of the presentation of distance in navigation applications, see Sect. 6.3.



(a) The audio guide reaches an intersection at a distance d in front of the user. The user still hears the guide straight ahead, and does not yet know if he or she should turn at the intersection

(b) The audio guide has made a turn to the right ahead of the user. At this point, the user hears that the audio guide has shifted somewhat to the right, and is aware of an upcoming turn to the right

(c) The user reaches the intersection and hears the audio guide straight to the right. Correspondingly, the user makes a turn to the right

Fig. 3 Illustration of the audio guide behaviour during route guidance

3.2.1 Route guidance

During route guidance, the audio guide moves at a certain distance ahead of the user, with its position continuously recalculated based on the position of the user and the geometry of the route. By listening to the location of the audio guide, the user thus becomes aware of an upcoming turn before reaching it.

The distance to the audio guide depends on the current speed of the user. During this study, the distance was set to be approximately 30 m for typical walking and 70 m for typical cycling speeds. Note that the user will not be able to detect an upcoming turn at this distance, but only once the audio guide has made the turn and moved a noticeable distance in the new direction. The behaviour of the audio guide is illustrated in Fig. 3.

The prototype software also takes into account that the user is not necessarily exactly on the route, but possibly some distance from it, and the audio guide is still placed in the direction that the user should go. For example, the user might walk on a sidewalk parallel to the route, in which case the guidance should not try to lead the user to the middle of the street, but along the street. This robustness of the audio guide placement also helps to avoid disturbing movements of the audio guide from side to side when there are inaccuracies in the GPS positioning, which is a typical phenomenon in city environments where high buildings are shadowing the GPS signals.

In the first of the three user studies (described in Sect. 4), the audio, in this case music, was presented in stereo mode when the user was walking along the correct route and there was no upcoming turn. In stereo mode, the music was heard as it would be with any normal music player. When the user approached an intersection, where he or she was supposed to make a turn, the music switched to guidance mode, and was thus heard as a point source from the direction of the audio guide (as illustrated in Fig. 3). Based on the results of this first user study (presented in Sect. 5.1.2) the switching between stereo mode and guidance mode was removed from the navigation service. Instead, the music was constantly presented in guidance mode in the two subsequent user studies.

3.2.2 Beacon guidance

In beacon guidance mode the audio guide is placed at the coordinates of the destination, and the task of finding a route to the destination is left to the user. The process of updating the sound direction thus depends only on the GPS position and the head orientation of the user. Far away from the destination, typical GPS inaccuracy does not cause large errors in the direction of the virtual sound source with respect to the user, but the sound rendering becomes more sensitive to these errors when the user gets closer to the destination. To make sure the user finds the target, it is therefore important that the arrival is registered at a suitable distance from the destination.

Table 1 Summary of user study conditions and participants

	User study A	User study B	User study C
Method	walking	walking	cycling
Location	city	city	suburbs
Tasks	route guid. 1 route guid. 2 beacon guid.	route guid. beacon guid.	route guid. beacon guid.
Counterbalanc.	no	yes	yes
Route guid.	guid. at turns	constant guid.	constant guid.
Hear-through	yes/no	yes	yes
Participants	12	10	12
– female/male	6/6	5/5	1/11
– age	21–41 (med. 31.5)	24–39 (med. 33.5)	22–43 (med. 29)

4 User studies

The user studies reported in this article were conducted to find answers especially to the following research questions:

- What is the overall user experience of using spatial music guidance for pedestrian and cyclist navigation?
- How well can people perceive the direction of spatialized music and use this direction to guide them to a destination?
- What are the preferences between route and beacon guidance?
- Do people feel safe when navigating with this type of music guidance?
- Can people pay attention to both the music guidance and the environment, and possibly have a conversation at the same time?
- Is spatialized mono music considered pleasant to listen to?
- What types of notifications do people want when using this type of navigation service and how do they want to control the service?
- How does the user experience differ when using hear-through headphones compared with headphones that block out sounds from the environment?

Three separate user studies were conducted and these are described in more detail below. The first study, user study A, was conducted with pedestrians in a city environment. Based on the results of this study, the navigation service was further developed, and a similar study, user study B, was conducted with the new version of the service. In user study C, the navigation service was evaluated with cyclists in a suburban environment. A summary of the user study conditions and participants is presented in Table 1.

The data collected during the user studies consisted of GPS location data of the routes that the participants took together with head orientation data. After each task, participants filled in a questionnaire regarding the task, and finally answered a number of questions verbally.

The familiarity of the user study environment varied between participants. Some of the participants were foreigners, and were not very familiar with the surroundings. The user studies were also conducted slightly away from the areas typically frequented by locals. None of the participants took part in more than one of the three user studies. The first user study took place in May, the second in June, and the third in September and October, 2014.

4.1 User study A

User study A consisted of three pedestrian navigation tasks. The first two tasks used route guidance, while the third used beacon guidance. All participants walked the same routes in the same order (although for the third task, using beacon guidance, the actual route taken by the participants varied, but the destination was the same). There was a short training task using route guidance before the first main task. In this user study, route guidance was implemented with switching between stereo mode and guidance mode, as described in Sect. 3.2.1.

The routes were chosen so that each of them took about 10 minutes to walk and included both left and right turns. The routes were located in the centre of Helsinki, but tall buildings along the routes were avoided to ensure adequate GPS signal reception. The routes were also tested to make sure that there were no strong magnetic disturbances affecting the head tracker. The chosen area was in general lightly-trafficked, but included some roads with slightly heavier traffic. Maps of the area together with the routes taken by the participants are shown in Fig. 22, 23, and 24 in Appendix A.

In this user study, we wanted to test if users prefer hearing sounds from the environment clearly while using the music guidance or having them attenuated, and how this would affect the navigation task. To achieve this, the hear-through microphone signals were attenuated completely during one of the first two tasks, and the environment was thus heard as attenuated by the in-ear headphones. During the other task the microphone signals were amplified so that the environment was heard at a normal and pleasant level. The order of these two conditions was counterbalanced between participants.

There were 12 participants in user study A: 6 female, 6 male, ages 21–41 (median 31.5). The participants were chosen to form a heterogeneous group of people working or studying in various fields. Five of the participants were students. All participants reported having normal hearing.

4.2 User study B

User study B consisted of two pedestrian navigation tasks, one using route guidance and one using beacon guidance. Two different routes were tested, with the second route starting where the first route ended. The order in which the two routes were walked together with the direction in which they were walked were counterbalanced between participants, as was the order of testing the two guidance types. In this user study, the guidance mode was constantly on during route guidance, as described in Sect. 3.2.1. There was no training task.

The routes for user study B were chosen from the same area as in user study A. Slightly longer routes were chosen, giving the participants approximately 15 minutes to test each guidance type. Otherwise the criteria for selecting the routes were the same as in user study A. Maps of the area together with the routes taken by the participants are shown in Fig. 25 and 26 in Appendix A. During this user study, the hear-through microphone signals were amplified so that the environment was heard at a normal and pleasant level.

There were 10 participants in user study B: 5 female, 5 male, ages 24–39 (median 33.5). The participants were chosen to form a heterogeneous group of people working or studying in various fields. Three of the participants were students. All participants reported having normal hearing.

One of the participants had problems with the left earpiece coming loose during the second task. This seriously affected the perceived direction of the music guidance and thereby the whole task, and for this reason the questionnaire answers given by this participant were removed from the final results.

4.3 User study C

User study C consisted of two cyclist navigation tasks, one using route guidance and one using beacon guidance. Two different routes were tested, with the second route starting where the first route ended. The order in which the two routes were cycled together with the direction in which they were cycled were counterbalanced between participants, as was the order of testing the two guidance types. In this user study, the guidance mode was constantly on during route guidance, as described in Sect. 3.2.1. There was no training task.

For safety reasons, a suburban area was chosen rather than a city centre for this study. The routes were chosen so that there were cycle tracks or lightly trafficked roads through residential areas for the participants to follow. The lengths of the routes were chosen so that they would take 15–20 minutes to cycle, with several turns both left and right. Maps of the area together with the routes taken by the participants are shown in Fig. 27, 28, 29, and 30 in Appendix A. During this user study, the hear-through microphone signals were amplified so that the environment was heard at a normal and pleasant level.

There were 12 participants in user study C: 1 female, 11 male, ages 22–43 (median 29). Participants who were familiar with listening to music while cycling, or thought that they would not have a problem with it, were recruited for this task. Most of the participants were working with technology-related research. Two of the participants were students. The third author took part in this user study as a participant. All participants reported having normal hearing.

4.4 Instruction of participants

Before taking part in the user study, participants were asked to sign a consent form. After this, they were given a background questionnaire to fill in. Before each task, the guidance method was explained. Participants were told that turning their head left and right typically helps in perceiving the direction of the guidance.

During user studies A and B, the experimenter occasionally had a discussion with the participants. This was done to investigate the cognitive load produced by following the music guidance, and whether the participants were able to focus on having a conversation in addition to following the guidance and paying attention to the environment. Participants were also encouraged to give comments about the navigation service during the tasks.

Participants were asked to choose one alternative from a small number of music artists available to listen to during each task. For each artist, a small selection of tracks were played. The artists represented musical genres such as pop, rock, jazz and hip hop. The participants were not able to adjust the volume of the music themselves, but could request the experimenter to do so at any time.

Participants were encouraged to walk or cycle at their own pace during the tasks. They were, however, asked to limit their speed if they usually cycled very fast, to prevent

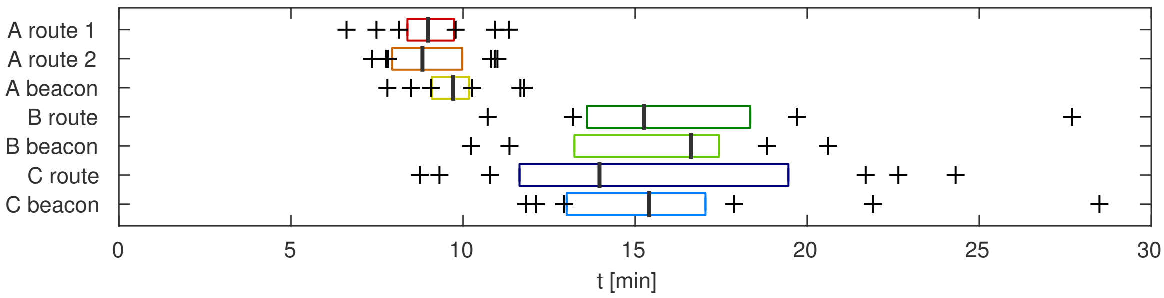


Fig. 4 Completion times for the different tasks

the distance between the participant and the experimenter from growing too large.

The participants were not told what the destination of the navigation tasks was. Neither were they allowed to look at the user interface of the navigation service or any other map. Before beacon guidance tasks participants were, however, told the approximate distance to the destination.

After each task, participants were given a questionnaire related to that task to fill in. The questionnaire mostly contained different statements; the participants specified how much they agreed or disagreed with these on a seven-point scale. After this, an oral interview took place. The tasks always ended close to a bench where the participant could sit down to fill in the questionnaire. Participants were rewarded with two movie tickets.

5 Results

Maps of the routes taken by the participants during the different user studies are shown in Appendix A. All participants completed the tasks successfully. In a few cases during route guidance, participants took a route that slightly deviated from the intended route, as shown on the maps. In some rare situations, when participants were apparently struggling with the task or, e.g., continued long past an intersection where they were supposed to turn, questions such as “from where do you hear the music right now?” were asked, partially to make sure that the difficulties were not related to technical problems. The participants were, however, not given any explicit advice on how to continue in these situations.

The completion times for the different tasks are shown as a box plot in Fig. 4. Note, that during user study A, both route guidance tasks as well as the beacon guidance task were conducted on different routes, meaning that these times are not comparable. During user studies B and C, two routes were used both for route and beacon guidance, with the order counterbalanced between participants. However, the exact route used for route guidance was not in all cases the shortest route between the starting point and the destination, since the routes were designed to have a suitable number of left and right turns.

5.1 Questionnaire and interview answers

Since the questionnaire answered by the participants contained a large number of questions, only the questions of most interest are presented here, in the form of box plots. Comments given by the participants during the interviews are summarized, and some individual comments are presented. The comments are not verbatim, as most of the interviews were conducted in Finnish.

A two-sided sign test was used to analyze the results, when appropriate. The number of pairs, S, where the data for the first condition were larger than those for the second condition is reported together with the p-value.

5.1.1 Overall impression

Participants in general liked the music guidance, with only a few participants slightly disagreeing (Fig. 5). When asked whether they would use this type of music guidance, participants said they would do so in most cases (Fig. 6). However, when cycling in a suburban area, as in user study C, participants would be more likely to use route guidance than beacon guidance ($S = 8, p = 0.039$). When walking in the city in user study B, there was no significant difference between the two guidance types ($S = 1, p = 0.38$).

5.1.2 Route vs. beacon guidance

The first route guidance task in user study A and the beacon guidance task in user study C were considered the most difficult tasks (Fig. 7). During user study A, there was a general, although not significant ($S = 2, p = 0.065$) preference for beacon guidance, as shown in Fig. 8. The main reason for this preference was presumably the fact that participants found the switching between stereo mode and guidance mode both disturbing and confusing. This problem was

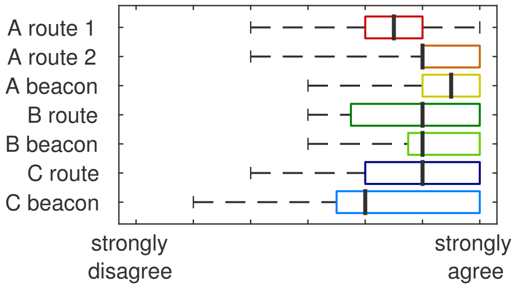


Fig. 5 "I liked the guidance a lot"

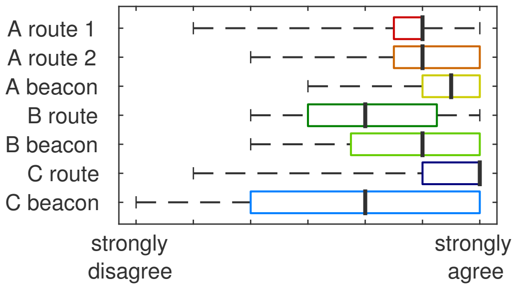


Fig. 6 "I would use this kind of music guidance"

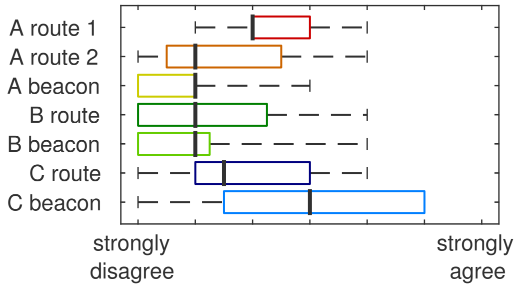


Fig. 7 "I felt the music guidance was difficult"

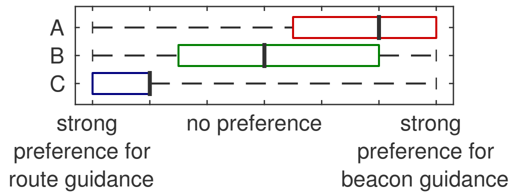


Fig. 8 "Did you prefer route or beacon guidance?"

emphasized by a clear increase in the loudness when switching to guidance mode, although this change also made it easier to notice an upcoming turn. The loudness difference was decreased after the first four–five tests, but many of the remaining participants still commented on the switching being annoying. The GPS location inaccuracy also made it impossible to switch to guidance mode at a same distance before each turn, which was considered confusing by many users.

Because of these findings, we decided to remove the switching between modes from the navigation service before user studies B and C, and guidance mode was instead constantly on in these user studies. This decision was also supported by the fact that few participants thought that the enjoyability of the music was considerably degraded by the spatialization in guidance mode.

During user study B, there was no general preference between route and beacon guidance ($S = 4, p = 1$). Instead, some participants preferred route guidance, others beacon guidance, while some liked and would use both guidance types. During user study C, most participants preferred route guidance ($S = 10, p = 0.012$), since choosing the route with beacon guidance proved difficult in the suburban environment compared with the city centre in user studies A and B.

Although there were some clear preferences for either guidance type in the user studies, many participants said during the interviews, that it depends largely on the context of the navigation, which guidance would work better. Many of the participants would more easily use beacon guidance if the environment was somewhat familiar, while route guidance was commonly considered better for unfamiliar environments. However, some participants explicitly said they would use beacon guidance even in a completely unfamiliar environment, and thought that it would be good for exploring new areas, e.g., as a tourist or on a leisurely bike trip. Some participants also commented that they rather would use route guidance if they were in a hurry, since it is supposed to guide you along the shortest or fastest route.

Some participants pointed out that a weakness of this type of route guidance is the lack of knowing when you should turn the next time before you approach the place where you are supposed to turn and the direction of the guidance begins to change. This fact was emphasized in these user studies, since participants did not know where their destination was, and could not predict where the guidance would lead them based on this. With beacon guidance, participants appreciated that they constantly were aware of the direction of the destination.

Several participants mentioned that beacon guidance might lead you into dead ends or otherwise problematic areas. Some participants suggested using some kind of combination of the two guidance types, where route guidance could be used when more specific guidance is needed. Participants also suggested having several successive beacons, i.e., waypoints, when guidance past obstacles is needed.

5.1.3 Knowing where to turn

There were differences in how easily participants could perceive the direction of the guidance (Fig. 9). For most participants it was quite easy, but a few participants seemed to struggle constantly with perceiving the direction. Many participants pointed out that head movements helped with the task.

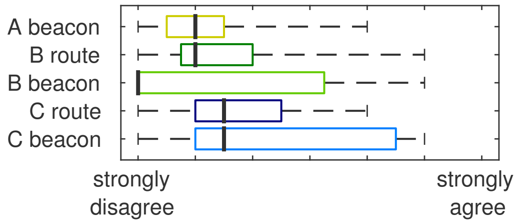


Fig. 9 "I think it was difficult to perceive the direction of the music guidance"

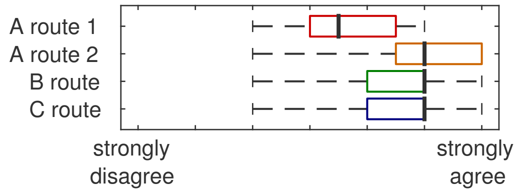


Fig. 10 "I think it was easy to notice when I was supposed to make a turn"

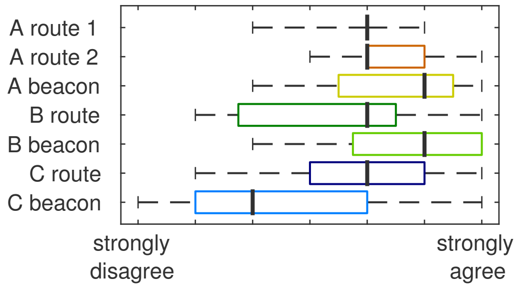
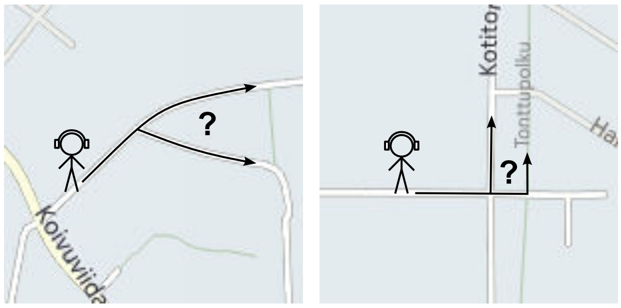


Fig. 11 "I think it was easy to know in what direction to continue at intersections"



(a) There are two alternative paths with less than a 90-degree angle in between. The guidance leads the user to the right, but the user cannot hear the direction precisely enough to be sure which path to take

(b) There are two intersections close to each other. If the guidance starts early, the user might turn at the first intersection, even though he or she is supposed to take the second path to the left

Fig. 12 Problematic situations during route guidance

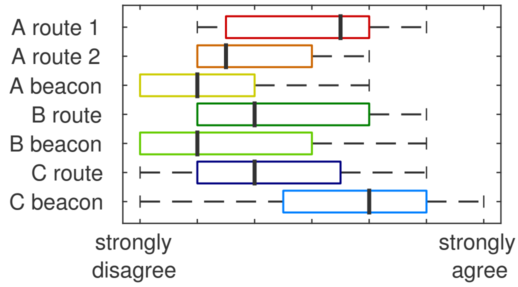


Fig. 13 "I was often uncertain whether I was going in the right direction"

With route guidance, participants in general thought that it was easy to know if they should make a turn at a right angle (Fig. 10 and 11). However, the exact direction of the guidance was generally considered difficult to perceive. Several participants commented that if there were two alternative paths, with less than a 90-degree angle in between, then it was difficult to know which path to take. This problem is illustrated in Fig. 12(a).

Participants generally thought that the music started guiding them in the right direction at the appropriate time before turns. However, participants mentioned that situations where there were two intersections close to each other (as illustrated in Fig. 12(b)) were confusing. If the guidance started leading the way before the participant reached the first intersection, he or she might make a turn there, even though he or she was supposed to turn at the second intersection.

Both of the problems illustrated in Fig. 12 might be emphasized in case the GPS location or map data are inaccurate. Problems similar to those shown in Fig. 12(b) have also been reported in another study [8].

In general, participants thought that the guidance mode stayed on slightly too long after turns during route guidance in user study A. Some participants commented that this made them feel unsure whether they were continuing in the right direction, since they waited for the music to go back to stereo mode to give them confirmation. Some participants also commented that prominent stereo effects in the music in stereo mode easily could be mistaken for the music switching to guidance mode.

In user studies A and B, which were conducted in the city centre, participants mostly thought it was easy to choose their route during beacon guidance (Fig. 11), since they usually only had to walk for a short while before they again could make a turn at the next intersection. The suburban surroundings in user study C proved more difficult. Participants commented that if they had to follow some path, which was not leading straight towards the destination, then they did not know how far they would need to cycle before they could make a turn towards the destination the next time. This is reflected in how confident participants were during the tasks (Fig. 13).

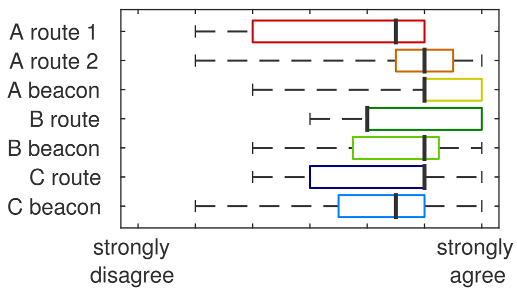


Fig. 14 "I think that it was easy to attend to the music guidance in the traffic"

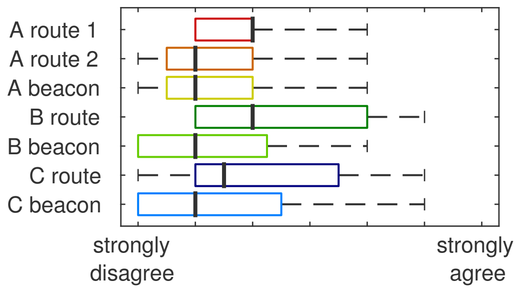


Fig. 15 "I think it was difficult to attend to the traffic and surroundings during the music guidance"

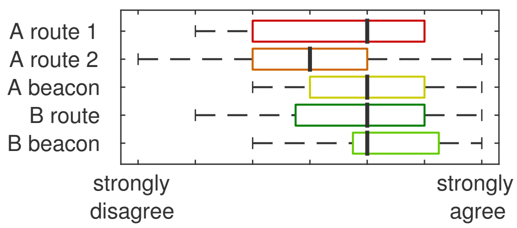


Fig. 16 "I think that chatting during the use of the music guidance was easy"

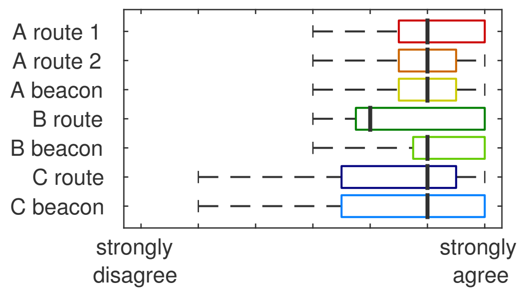


Fig. 17 "I felt safe while using the guidance"

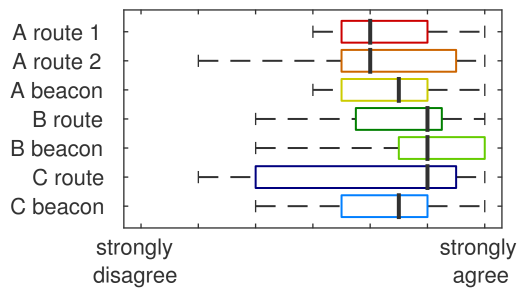


Fig. 18 "I could hear my surroundings clearly during the task"

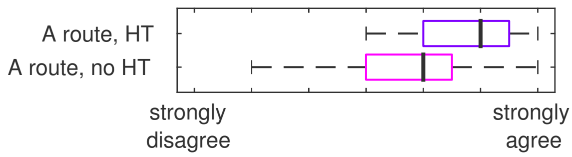


Fig. 19 "I could hear my surroundings clearly during the task." Answers for the two different conditions during route guidance in user study A. During the *HT* (hear-through) condition, the hear-through functionality of the headphones was enabled, so that the environment was heard at a natural level. During the *no HT* condition, the hear-through functionality was disabled

5.1.4 Attention

In general, participants said that they were able to pay attention to both the music guidance and the environment at the same time (Fig. 14 and 15), although some participants found this more difficult, possibly depending on the volume of the music and the amplification of the hear-through microphone signals. Having a conversation with the experimenter at the same time was generally considered more challenging (Fig. 16), although some participants thought it was easy. This was presumably partially due to the cognitive load caused by navigation in the city, and partially due to the music masking the voice of the experimenter. Many participants pointed out that they would not use this type of navigation service in the company of others.

5.1.5 Sense of safety

Most participants did not feel that using music for guidance reduced their sense of safety compared with listening to music in general when walking or cycling, and overall felt safe while using the navigation service (Fig. 17). Many of them pointed out that it felt safer than navigating by looking at a map, since they constantly could keep their eyes on the surroundings. Some participants, however, noticed that they focused on the guidance at intersections at the expense of paying attention to the traffic.

5.1.6 Hear-through

Participants reported that they in general could hear the environment quite well during the tasks (Fig. 18). When the hear-through functionality of the headphones was disabled, sounds from the environment were blocked out to a certain degree, but not very effectively (Fig. 19). A sign test revealed no significant difference between the two conditions ($S = 6, p = 0.29$).

When listening to music on the go, some of the participants said that they want to hear the environment clearly, while others want to concentrate on the music and block

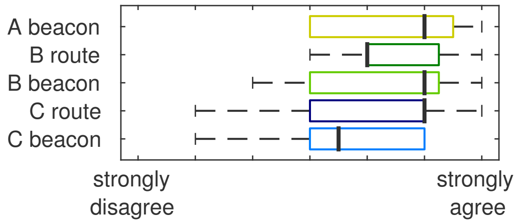


Fig. 20 "I think that listening to the music was pleasant compared with normal stereo listening"

out sounds from the environment. Most participants said that they heard the environment naturally when the microphone signals of the hear-through headset were appropriately amplified.

5.1.7 Impact of the music

Participants pointed out that some artists and tracks, especially those with more dynamics, were more difficult to use for guidance than others. Quiet parts and transitions between tracks were perceived as particularly problematic.

Some participants mentioned that being able to listen to music that they liked had a strong effect on how they felt during the tests. Even if they thought that they were making some bad route choices with beacon guidance, and even felt totally lost at some points, they felt relaxed and enjoyed the experience. Similar observations were done in the evaluation of Melodious Walkabout [3].

5.1.8 Mono vs. stereo music

During route guidance in user study A, when the music was presented in stereo mode, participants heard the music as with any normal music player (as described in Sect. 3.2.1). In all other cases during the three user studies, the music was mixed to mono and played from the direction of the audio guide as a point source using HRTFs. Most participants did not find it particularly disturbing that the music was not played in normal stereo (Fig. 20). Similar observations where done in the study by Fujimoto and Turk, where only a minority of the participants thought that spatialization of the music would detract considerably from their enjoyment of it [5].

Many participants pointed out that they heard the music equally in both ears when they were walking in the right direction. After the beacon guidance task in user study B, one participant commented that "if I stay on the route, then the music is always in stereo." However, some participants found it slightly disturbing if they were unable to walk straight towards the destination during beacon guidance, and the music consequently was heard from either side.

5.1.9 Headphones

How well the headphones suited the participants varied greatly. In general, participants seemed to have more problems when cycling with the headphones without a headband in user study C, since these did not stay well in the ears of some of the participants. The wind noise that was heard through the hear-through headset was also found disturbing by several participants, and the microphone-signal amplification was therefore decreased in a few cases.

Many participants said that they prefer small unnoticeable in-ear headphones for mobile usage. Other participants would rather use large headphones with a headband, but some participants pointed out that these are difficult to store, as they do not fit in a pocket, and that they might be problematic, e.g., in combination with certain hairstyles.

Wireless headphones were suggested by a couple of participants, since the wires to the headphones in the test setup were too short and thereby restricted large head movements. Finally, a couple of participants pointed out that they would not want to buy headphones separately to use with this type of service, but suggested that people should be able to use any headphones they want to.

5.1.10 Notifications and controlling the service

Most participants thought that the female voice that informed them that they had successfully completed the navigation task, by saying "you have reached your destination," suited its purpose well. In general, speech was deemed good for such notifications, because it is clear and stands out from the music.

When asked about other notifications they would want, many participants mentioned that it would be useful to know the distance to the destination. Participants suggested notifications both halfway through the route and when approaching the destination (e.g. a hundred meters away) as well as notifications available at the push of a button. A couple of participants would have wanted a notification about the remaining distance once in a while, ideally at adjustable intervals.

Several participants also suggested that the navigation service could provide information about nearby points of interest, e.g., places to eat. Other suggestions included:

– "It is not necessary to have other notification sounds with this type of route, but if there is, e.g., a large intersection, a wide street to cross, or many possible streets at one intersection, then additional information would be useful."
– "It could inform that you have continued in the wrong direction at an intersection, either with speech or with changing timbre etc."
– "If you are walking completely in the wrong direction, you could get a notification. Speech is good for this."

- "Beacon sounds from distinct landmarks would help you to perceive your location in the city."
- "If there was a dead end ahead, or a long straight without any possibility to make a turn, then the service could warn about it."
- "There could be some other kind of sound effect to inform you that you should, e.g., cross the road."

Some participants, however, commented that they would want as little notification sounds as possible, as they want to focus on listening to the music. Many of them also commented that a large part of the attractiveness of the system is that no extra sounds are present, only the music.

Some participants did not feel the need to control the service during the tests, while others suggested different features they would want to control. The most common things mentioned were choosing or skipping tracks, adjusting the volume of the music, and adjusting how well the environment was heard. Automatic adjustment of the music volume and the amplification of the hear-through microphone signals was also suggested. Participants also wanted to be able to personalize notifications, and control when the guidance starts before an intersection. Other suggestions included:

- "I would like to be able to ask for directions in unclear situations."
- "Turning the beacon on and off with a button, so that you can listen to normal stereo when you want to."
- "You might want to add a waypoint to the route."

Many participants thought it would be fine to take the phone out of the pocket to control something, while others suggested buttons on the headphones or the cable, or using voice commands. Several participants, however, had bad experiences with voice commands or were otherwise sceptical about using these.

5.1.11 The need to see a map

The wish to see a map varied from participant to participant (Fig. 21). Some participants did not even consider looking at a map in the test situation, while others would have wanted to have a map. Many of the participants would in a real navigation situation look at a map at least once in the beginning, to get an overview of the route and the area. When cycling using beacon guidance in the suburban environment in user study C, participants had a much stronger urge to see a map than during the other tasks.

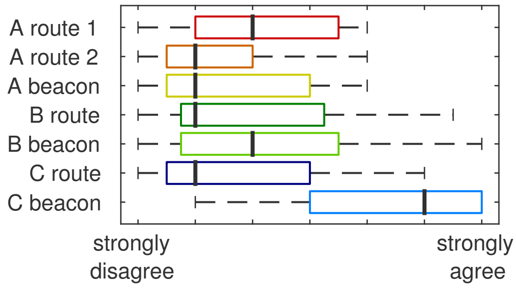


Fig. 21 "I would often have liked to see the route on a map"

5.2 Order effects

When comparing the questionnaire answers given in user study A after the first route guidance task with those given after the second task, some noticeable differences can be seen. In particular, a sign test reveals significant differences between the two tasks for the following questions:

- "I felt the music guidance was confusing." ($S = 8, p = 0.008$)
- "I felt the music guidance was clear." ($S = 1, p = 0.039$)
- "I liked the guidance a lot." ($S = 0, p = 0.031$)
- "It was easy to trust the guidance." ($S = 0, p = 0.008$)
- "I think it was easy to notice when I was supposed to make a turn." ($S = 0, p = 0.002$)
- "I think it was easy to know in what direction to continue at intersections." ($S = 0, p = 0.016$)
- "I think that the music used in the guidance fit this purpose well." ($S = 0, p = 0.031$)

In all cases the answers were more positive for the second task. These differences may be attributed to several factors. It is likely that there was a learning effect involved, which was emphasized in the results because of the short lengths of the routes in user study A. After the second task, some participants commented that they felt that they now had got used to the navigation service. The differences may, partially, also be related to the different routes where these tasks were performed; the first route contained areas with roadworks and more traffic than the second route.

6 Discussion and design recommendations

This section discusses the results of the user studies and based on these presents recommendations for designing a navigation service using music guidance for pedestrians or cyclists.

6.1 Limitations of the current study

The main limitation of the conducted user studies is that they did not represent entirely realistic use cases. In a realistic situation, the user would know his or her destination, and would have the option to look at a map beforehand, to have a general idea of the route to the destination. In these user studies, however, participants were not told what the destination was, as some of them were familiar with the area and would have found their own way based on this knowledge. Neither were they shown any kind of map, as the idea was to test how they could navigate to the destination with the help of the music guidance alone. Thus, performing naturalistic user studies would provide further insight into natural use of a music guidance service.

6.2 Guidance

In a navigation service with music guidance, and probably other types of spatial audio guidance as well, both route guidance and beacon guidance should ideally be provided as alternatives, as both of these have their advantages and disadvantages, and their suitability depends on the environment and the user's preferences. Switching between guidance modes should be easy to do during navigation, so that, e.g., a user may switch from beacon guidance to route guidance if he or she feels unsure of how to continue.

Practical use of music guidance, especially beacon guidance, would presumably include looking at a map at least once before heading off and possibly one or several times at later stages. An overview map of the route, or area between the starting point or current location and the destination in case of beacon guidance, should thus be readily available to the user.

In these user studies, participants were not allowed to look at a map, as the intention was to test how well they could find their way by just following the music and without utilizing any prior knowledge about the area to a large extent. Compared with the results of this study, the preference between the two guidance types might thus be slightly more positive towards beacon guidance in practical situations, especially when cycling.

During the cycling tasks in the suburban environment in user study C, participants felt unsure about finding their way with beacon guidance, and therefore generally preferred route guidance. However, most of the participants still found a good route to the destination using beacon guidance (as can be seen by comparing the routes in Fig. 27, 28, 29, and 30, as well as comparing the completion times in Fig. 4). The preference for route guidance was thus rather related to the feeling of uncertainty during beacon guidance than to the actual performance in finding the way to the destination.

6.2.1 Route guidance

If there are alternative paths at less than 90-degree angles (as illustrated in Fig. 12(a)), more exact guidance is needed. In these user studies, participants felt that they could not perceive the direction of the music guidance precisely enough in these situations. If the precision and accuracy of the music guidance cannot be improved (e.g. by decreasing the sound source width), further guidance should be made available to users, e.g., by the use of verbal instructions in ambiguous situations. In some cases, when there is no risk of confusion, the guidance could be exaggerated in the appropriate direction to make it clearer where the user should turn.

If, for example, the user is supposed to turn left, and there are several paths leading to the left close to each other (as illustrated in Fig. 12(b)), the guidance should take this into account, as well as possible, so as not to guide the user on to a path prior to the correct path. This becomes more problematic if, e.g., the user is cycling at a high speed. In this case, verbal instructions could be used to provide additional information.

If the user makes a wrong turn and strays from the intended route, clear and immediate feedback should be given. However, if a sensible rerouting alternative is available, a new route should instead be suggested, allowing the user to choose his or her own route, with the guidance only providing suggestions on how to continue towards the destination. Brown and Laurier [27] observed drivers using turn-by-turn guidance and found it common not to follow the suggested route for various reasons. They suggested that users should not be forced back on a route that they clearly do not want to follow. If the user repeatedly ignores instructions, this should instead be taken as an indication that the user wants to take another route.

If the map data used by the navigation service provide detailed information about which side of the street cycle tracks and sidewalks lie on, and where these continue, this should be used to direct the user to the right side of the street, if the path on the other side of the street continues in the wrong direction. Although obtrusive, speech is probably the best means to convey this type of information. Spatial audio cannot be used for this purpose, since the GPS location is not accurate enough to infer which side of the street the user currently is on.

Constant guidance was in these user studies more successful than guidance only during turns. Participants found the switching between guidance modes in user study A disturbing, especially because of the large loudness difference between stereo mode and guidance mode during the first four–five tests. In addition, despite the training task before the actual tests, an order, and thus, probably, learning effect could be observed in this user study. Some participants also commented that they were getting used to the guidance only after the second task. However, in user study B, where there was constant guidance, participants in general seemed to "get the idea" much quicker and without a training task. Most participants also found the music enjoyable even though it was not presented in normal stereo format, which suggests that the separate stereo mode in the first user study was unnecessary.

6.2.2 Beacon guidance

Beacon guidance should support waypoints, to give more accurate guidance past obstacles when needed. The user could choose to add waypoints either automatically or manually. As Jones et al. [1] point out, such waypoints should have a capture radius that depends on the location and the proximity of other waypoints. For example, a waypoint in a large open field should have a large capture radius, while waypoints in a narrow passageway should have a small capture radius.

Users should be given the option to get advice about their choice of route, either at crossroads or after they have chosen a route that leads to a dead end. Advice or reassurance during long straights where the user cannot turn towards the beacon would also be helpful.

6.2.3 Perceiving the direction of the guidance

For a small number of participants, perceiving the direction of the virtual music source was considerably more difficult than for others. The reasons for this are unclear, but several possible explanations exist. It may, for example, be that the HRTFs used in the navigation service suited these participants poorly. Another possible explanation is related to the problems that one participant in user study B had: one of the earpieces was not sitting well in her ear, and therefore the sound from that earpiece was attenuated considerably. It may be that other participants also had problems related to how the earpieces fit in their ears, which might have led to deterioration of localization cues, particularly the interaural level difference.

6.2.4 Music vs. other audio for guidance

Participants suggested, e.g., radio programmes and audio books as alternatives to music for this kind of guidance. Compared with music, such audio sources might, however, be problematic, as some participants pointed out. Music is often, but not in all cases, a continuous stream of sound, where noticeable pauses typically occur only when the track changes. Speech programmes, on the other hand, contain frequent short pauses in the stream of sound. Although these pauses seldom are very long, even the short pauses when one track changed to another during these user studies were considered by the participants to be problematic. Nevertheless, users should ideally be given the option to use any audio source of their choosing to be used for the guidance. For example, they might already be listening to music using their favourite music player and just enable the guidance when needed.

Although this conflicts with the possibility to use any audio player, music tracks should, if possible, be crossfaded so that there are no silent parts between the tracks. Silent parts during the tracks might be equally problematic, and are more difficult to address. The navigation service might fill such silent parts with some easily localizable and not too disturbing sound, either automatically or, e.g., by the press of a button.

6.3 Presentation of distance

Using the volume of the music as a representation of the distance to the destination or next waypoint is an intuitive mapping that has previously been suggested and to some extent also tested [1–5]. However, no extensive evaluation of this approach has been reported.

Intuitive though it may seem, this approach is in many ways problematic. Music listeners tend to adjust the volume of the music to a pleasant level, which depends on several different factors. These include the type of music, the level of background noise, the presence of other sounds that the user wants to hear, the type of headphones used, and personal preferences. Although these conditions probably leave a range of appropriate levels that could be utilized, this range is presumably not very large. Thus, only a very crude representation of distance will be available by mapping it to this range of levels.

In addition, as this range of appropriate levels varies, music volume can only be used as a relative cue. The user will know if he or she is walking towards or away from the destination as the volume changes, but will not know with any reasonable precision how long a distance still remains. In their experiments, Fujimoto and Turk gave a reference sample representing a certain distance between each sample that was interpreted by the participants [5], but providing a reference sample at regular intervals does not seem like an enjoyable approach in real applications. Further, dynamic changes in the music tracks themselves and loudness differences between tracks may be mistaken for distance cues or mask the actual cues.

Liljedahl and Lindberg [28] suggested and evaluated alternative sound parameters for expressing distance on a relative scale. These parameters were pitch, reverberation, low-pass filtering, and combinations thereof. Reverberation was shown to be the most reliable parameter, while pitch was found to be an unreliable and ambiguous cue to distance. Reverberation could thus be a potential parameter for conveying the distance to the destination in a navigation application. However, excessive addition of reverberation might affect the enjoyability of the music, and reverberation already present in the music might also affect the interpretation of distance cues.

Adding distance cues to the music can thus not be recommended without reservation. However, many of the user study participants would have wanted information about the distance to the destination. Based on this, using speech to present the distance when desired seems a viable option, as most participants felt that speech was a good method for conveying necessary information. Participants suggested notifications about the remaining distance to the destination when halfway through the route, when getting close to the destination, at adjustable intervals, or by the push of a button. Alternatively, the distance could be presented using, e.g., earcons, spearcons, or short pulses [29], but in those cases users must learn the mapping between distance and these sounds.

6.4 Safety

The participants mostly found it easy to pay attention to both the music guidance and the environment at the same time, suggesting that the cognitive load caused by the guidance is relatively low. In user studies A and B, the third simultaneous, partially auditory task of having a conversation was considered slightly difficult, but this is at least to some extent related to the auditory masking caused by the music. In general, participants thought that using the navigation service did not decrease their sense of safety compared with normal music listening while walking or cycling.

Based on these observations and previous studies on pedestrian and cyclist safety (presented in Sect. 2.5), we believe that a well-designed navigation service using music guidance does not pose other safety risks than the environmental isolation caused by listening to music using headphones. As this type of service presumably would be used only by people who otherwise also listen to music, it might actually reduce safety risks, by removing inattentional blindness caused by looking at maps or navigation instructions on a mobile phone screen while walking or cycling.

The guidance given at intersections in a navigation service should be designed and fine-tuned with safety in mind. Unclear guidance or drastic changes to the guidance given when the user already is in an intersection might draw attention away from the surroundings and thereby decrease safety.

6.5 Headphones

Some people like to be able to hear their environment when listening to music while walking or cycling, while others want to block the environment out and only hear the music. Microphone-hear-through headphones allow both, and allow users to change how much they hear of the environment depending on the situation. As an option, the hear-through level could automatically be adjusted based on the context.

6.6 Notifications

Notifications during the guidance should be kept at a minimum. Necessary and user-enabled notifications should be presented with speech, since the meaning of the notifications is clear in that case, and they stand out from the environment, as pointed out by several participants. Information about nearby points of interest could be presented through the navigation service, if the user desires.

6.7 Walking vs. cycling

Based on these user studies, both guidance types should be equally usable for both walking and cycling. The different preferences for beacon and route guidance in user studies B and C are presumably to a large degree due to the different environments in which the tasks were performed, rather than due to the method of transportation.

When cycling, speeds are typically much higher than when walking, so guidance at intersections should be provided earlier so that the user has time to react. This is already taken into account in the implementation of our navigation service, where the distance to the audio guide depends on the current speed of the user. However, more exact experiments regarding the optimal parameters determining this distance are still needed, so that the same parameters are usable for walking, running, and cycling at various speeds. Users should also be able to adjust these parameters to their own liking. Another difference to consider is that when cycling, the user has more limited possibilities to turn his or her head to perceive the exact direction of the guidance.

7 Conclusions

In this article, the results from three user studies on using spatial music guidance for pedestrian and cyclist navigation were presented. In these studies, two different guidance methods, route and beacon guidance, were tested. The answers and comments from the participants clearly indicate that both guidance methods are valid approaches, but also that they are suitable for different situations. The personal preference between these two methods also varied considerably from participant to participant. In general, beacon guidance was seen as useful in more familiar surroundings, but also, somewhat contradictory, for touristic exploration. Route guidance, on the other, was generally seen as a better alternative when the user is in a hurry or is not very familiar with the surroundings.

While many previous studies on music guidance for navigation have adjusted the stereo balance of the music to indicate direction, we used HRTFs to present both channels of the music from a single direction. In general, participants did not feel that this considerably affected the enjoyability of the music, and most participants felt it easy to perceive the direction of the music presented this way. As previous studies using stereo balance adjustment also have shown promising results, this suggests that both HRTFs and stereo balance adjustment are viable approaches. Further studies comparing the two techniques would be needed to pinpoint the practical advantages and disadvantages of using either method.

The results of the three user studies presented in this article show that using a spatialized music source for guidance is a promising approach for pedestrian and cyclist navigation services. In general, participants in these user studies found the music guidance easy to follow, liked the guidance, and would use this type of guidance as well as recommend it to others.

Acknowledgements We would like to thank all participants in the user studies for their feedback and suggestions, as well as David McGookin for comments and discussion.

References

1. Jones M, Jones S, Bradley G, Warren N, Bainbridge D, Holmes G (2008) ONTRACK: Dynamically adapting music playback to support navigation. Pers Ubiquit Comput 12(7):513–525, doi:10.1007/s00779-007-0155-2
2. Strachan S, Eslambolchilar P, Murray-Smith R, Hughes S, O'Modhrain S (2005) GpsTunes: Controlling navigation via audio feedback. In: Proc MobileHCI '05, pp. 275–278, Salzburg, Austria, doi:10.1145/1085777.1085831
3. Etter R, Specht M (2005) Melodious Walkabout: Implicit navigation with contextualized personal audio contents. In: Adjun Proc 3rd Int Conf Pervasive Comput, pp. 43–49, Munich, Germany
4. Zwinderman M, Zavialova T, Tetteroo D, Lehouck P (2011) Oh music, where art thou? In: Proc MobileHCI '11, pp. 533–538, Stockholm, Sweden, doi:10.1145/2037373.2037456
5. Fujimoto E, Turk M (2014) Non-visual navigation using combined audio music and haptic cues. In: Proc 16th Int Conf Multimodal Interact, pp. 411–418, Istanbul, Turkey, doi:10.1145/2663204.2663243
6. Calvo A, Finomore V, Burnett G, McNitt T (2013) Evaluation of a mobile application for multimodal land navigation. Proc Hum Fact Ergon Soc Annu Meet 57(1):1997–2001, doi:10.1177/1541931213571446
7. Calvo A, Finomore V, McNitt T, Burnett G (2014) Demonstration and evaluation of an eyes-free mobile navigation system. Proc Hum Fact Ergon Soc Annu Meet 58(1):1238–1241, doi:10.1177/1541931214581258
8. Liljedahl M, Lindberg S, Delsing K, Polojärvi M, Saloranta T, Alakärppä I (2012) Testing two tools for multimodal navigation. Adv Hum Comput Interact 2012, doi:10.1155/2012/251384, article ID 251384
9. Väänänen R, Vesa S, Hämäläinen M (2014) Testing the user experience of an augmented reality headset and 3D audio-guided pedestrian navigation. In: 55th Int Audio Eng Soc Conf, Helsinki, Finland
10. Blauert J (1997) Spatial hearing: The psychophysics of human sound localization. MIT Press, Cambridge, MA, USA
11. Begault DR, Wenzel EM (1992) Techniques and applications for binaural sound manipulation in human-machine interfaces. Int J Aviat Psychol 2(1):1–22, doi:10.1207/s15327108ijap0201_1
12. Begault DR, Wenzel EM, Anderson MR (2001) Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. J Audio Eng Soc 49(10):904–916
13. Klatzky RL, Marston JR, Giudice NA, Golledge RG, Loomis JM (2006) Cognitive load of navigating without vision when guided by virtual sound versus spatial language. J Exp Psychol Appl 12(4):223–232, doi:10.1037/1076-898X.12.4.223
14. Lokki T, Gröhn M, Savioja L, Takala T (2000) A case study of auditory navigation in virtual acoustic environments. In: Proc 6th Int Conf Auditory Display, pp. 145–150, Atlanta, GA, USA
15. Larsen CH, Lauritsen DS, Larsen JJ, Pilgaard M, Madsen JB (2013) Differences in human audio localization performance between a HRTF- and a non-HRTF audio system. In: Proc 8th Audio Mostly Conf, Piteå, Sweden, doi:10.1145/2544114.2544118
16. Gonot A, Chateau N, Emerit M (2006) Usability of 3D-sound for navigation in a constrained virtual environment. In: Audio Eng Soc Conv 120, Paris, France, paper no. 6800.
17. Tran TV, Letowski T, Abouchacra KS (2000) Evaluation of acoustic beacon characteristics for navigation tasks. Ergonomics 43(6):807–827, doi:10.1080/001401300404760
18. Tikander M, Karjalainen M, Riikonen V (2008) An augmented reality audio headset. In: Proc 11th Int Conf Digit Audio Effects, pp. 181–184, Espoo, Finland
19. Pielot M, Poppinga B, Heuten W, Boll S (2012) Tacticycle: Supporting exploratory bicycle trips. In: Proc MobileHCI '12, pp. 369–378, San Francisco, CA, USA, doi:10.1145/2371574.2371631
20. Lichenstein R, Smith DC, Ambrose JL, Moody LA (2012) Headphone use and pedestrian injury and death in the united states: 2004–2011. Inj Prev 18(5):287–290, doi:10.1136/injuryprev-2011-040161

21. Jensen BS, Skov MB, Thiruravichandran N (2010) Studying driver attention and behaviour for three configurations of GPS navigation in real traffic driving. In: Proc CHI '10, pp. 1271–1280, Atlanta, GA, USA, doi: 10.1145/1753326.1753517
22. Thompson LL, Rivara FP, Ayyagari RC, Ebel BE (2013) Impact of social and technological distraction on pedestrian crossing behaviour: an observational study. Inj Prev 19(4):232–237, doi:10.1136/injuryprev-2012-040601
23. Walker EJ, Lanthier SN, Risko EF, Kingstone A (2012) The effects of personal music devices on pedestrian behaviour. Saf Sci 50(1):123–128, doi: 10.1016/j.ssci.2011.07.011
24. Nasar J, Hecht P, Wener R (2008) Mobile telephones, distracted attention, and pedestrian safety. Accid Anal Prev 40(1):69–75, doi:10.1016/j.aap.2007.04.005
25. de Waard D, Schepers P, Ormel W, Brookhuis K (2010) Mobile phone use while cycling: Incidence and effects on behaviour and safety. Ergonomics 53(1):30–42, doi: 10.1080/00140130903381180
26. Wenzel EM (2001) Effect of increasing system latency on localization of virtual sounds with short and long duration. In: Proc 7th Int Conf Auditory Display, pp. 185–190, Espoo, Finland
27. Brown B, Laurier E (2012) The normal, natural troubles of driving with GPS. In: Proc CHI '12, pp. 1621–1630, Austin, TX, USA, doi:10.1145/2207676.2208285
28. Liljedahl M, Lindberg S (2011) Sound parameters for expressing geographic distance in a mobile navigation application. In: Proc 6th Audio Mostly Conf, pp. 1–7, Coimbra, Portugal, doi:10.1145/2095667.2095668
29. Hussain I, Chen L, Mirza HT, Xing K, Chen G (2014) A comparative study of sonification methods to represent distance and forward-direction in pedestrian navigation. Int J Hum Comput Interact 30(9):740–751, doi: 10.1080/10447318.2014.925381

A Maps of the routes

The routes taken by participants during route guidance in user study A are shown in Fig. 23 and 24. In this study, route guidance was tested during two tasks, with the second task starting where the first task ended, after the participants had answered the questionnaire and interview questions. The beacon guidance task started where the second route guidance task ended. The routes taken by the participants during beacon guidance are shown in Fig. 22. During this user study, the three tasks were always performed in the same order with the routes walked in the same direction.

The routes taken during route guidance in user study B are shown in Fig. 25 and the routes taken during beacon guidance are shown in Fig. 26. Two routes (starting points and destinations) where used in this study, with the second route starting where the first route ended. The order of the tasks (route and beacon guidance) and the order and direction of the two routes were counterbalanced between participants.

In user study C, two routes were again used, with the second route starting where the first route ended. The order of the tasks (route and beacon guidance) was counterbalanced between participants, as was the order and direction of the two routes. The routes that participants cycled using route guidance are shown in Fig. 27 and 29. The routes that participants cycled using beacon guidance are shown in Fig. 28 and 30.

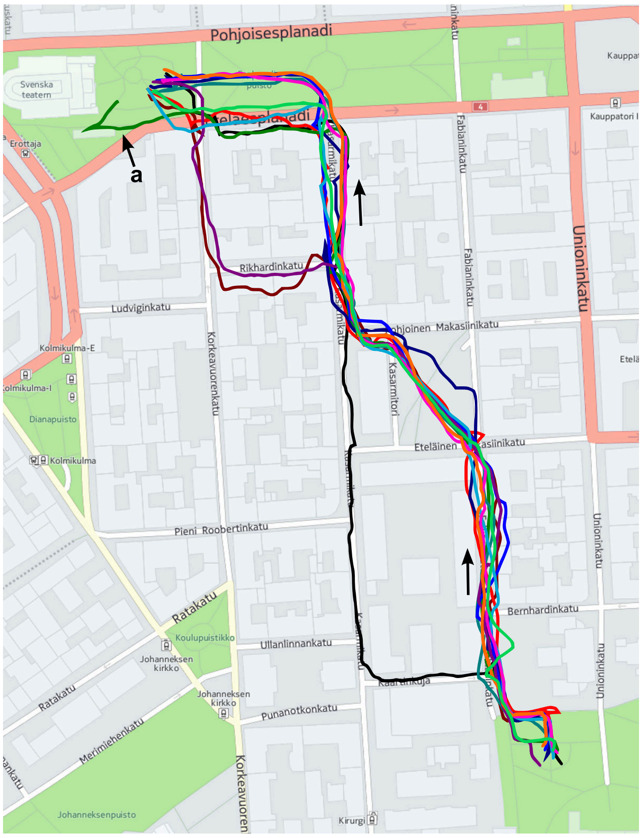


Fig. 22 The routes taken by the participants during beacon guidance in user study A. If no indication is given as to how long the remaining distance to the destination is, participants might make suboptimal route choices especially towards the end of the route, as one participant did at point *a*

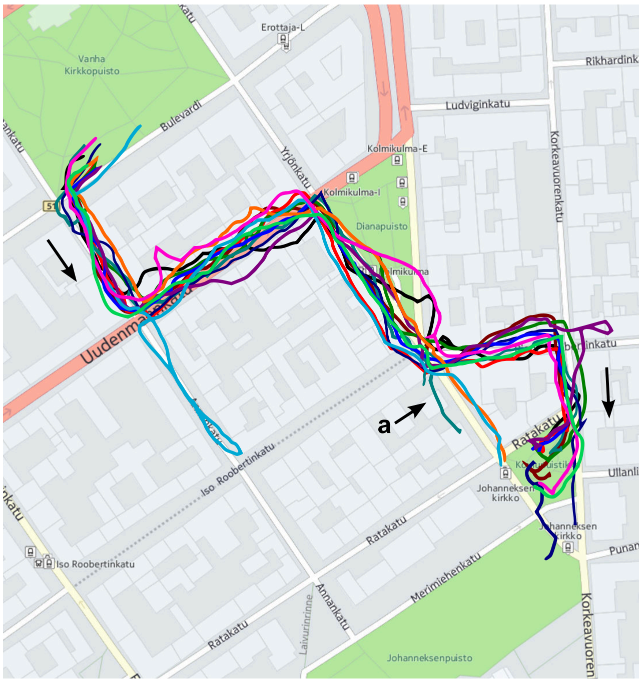


Fig. 23 The routes taken by the participants during the first route guidance task in user study A. Turns that were not at right angles proved difficult, as was the case at point *a*. There, three participants continued on straight, when they were supposed to make a 60° turn to the left

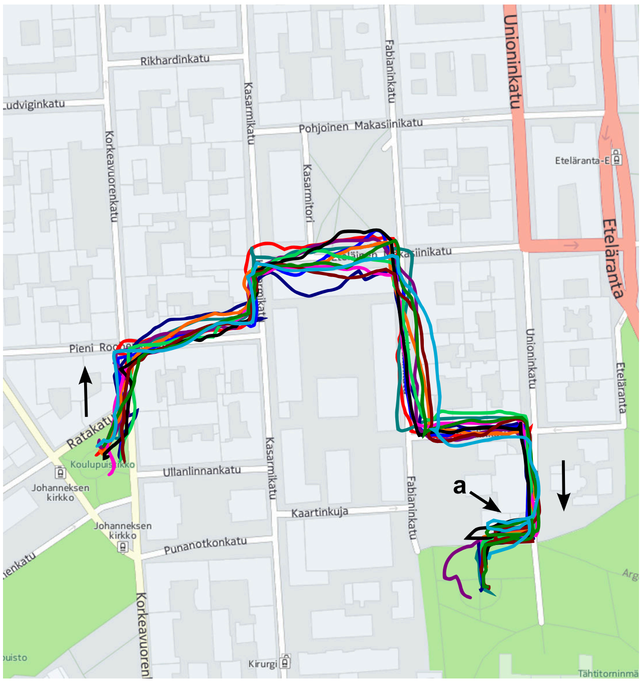


Fig. 24 The routes taken by the participants during the second route guidance task in user study A. At point *a*, there was a possibility to make a right turn into a schoolyard just before the right turn that the participants were supposed to make. Four of the participants walked into the schoolyard and back, while some other participants said that they guessed that they were not supposed to go into the schoolyard

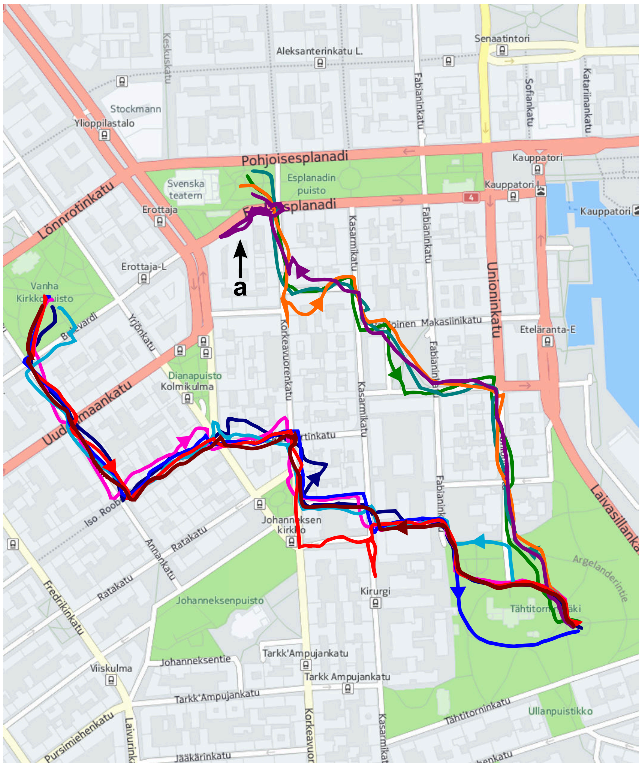


Fig. 25 The routes taken by the participants during route guidance in user study B. Turns that were smaller than 90 degrees caused confusion, for example at point *a*, where one participant had difficulties realizing that she was supposed to make a 45-degree turn at the intersection

Fig. 26 The routes taken by the participants during beacon guidance in user study B. The odd route taken by one participant at point *a* is due to the left earpiece coming loose, so that the participant heard sound mostly from the right earpiece. At point *b*, the long straight without any possibility to turn towards the destination caused some uncertainty among participants

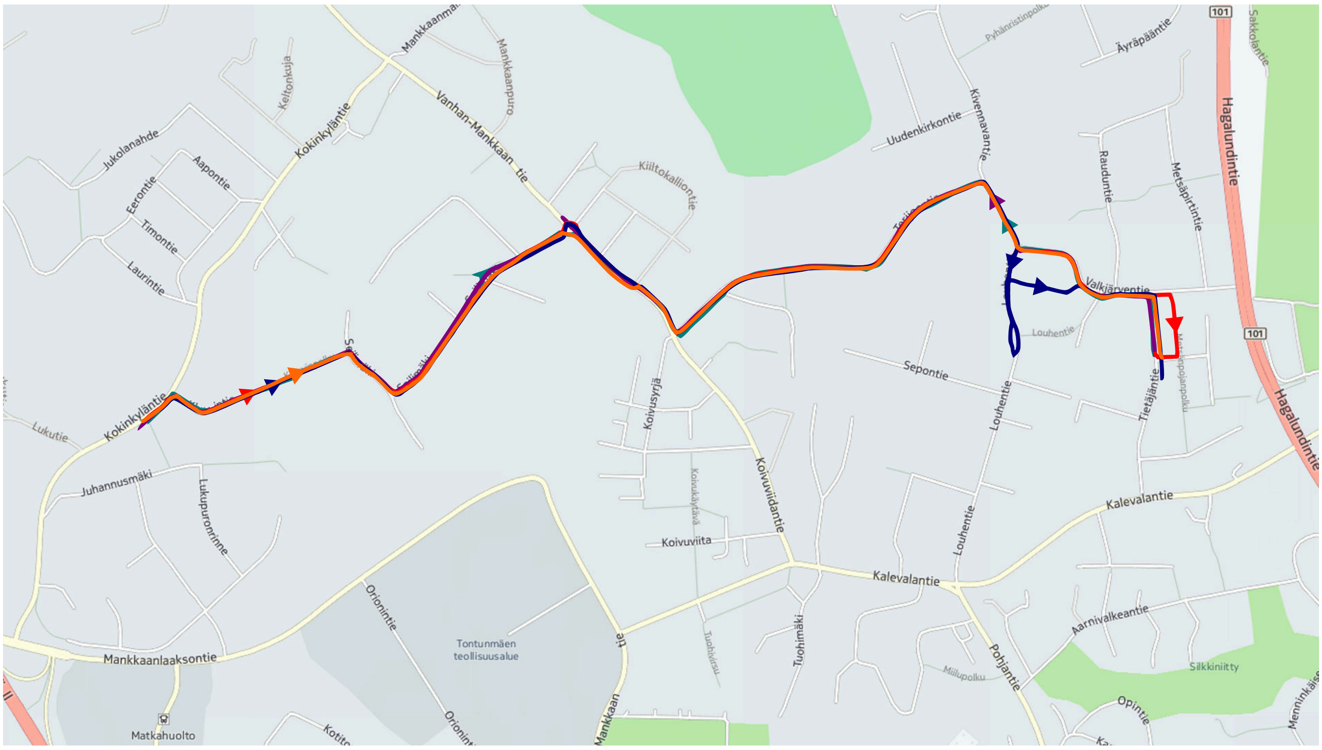


Fig. 27 The northern routes taken by the participants during route guidance in user study C

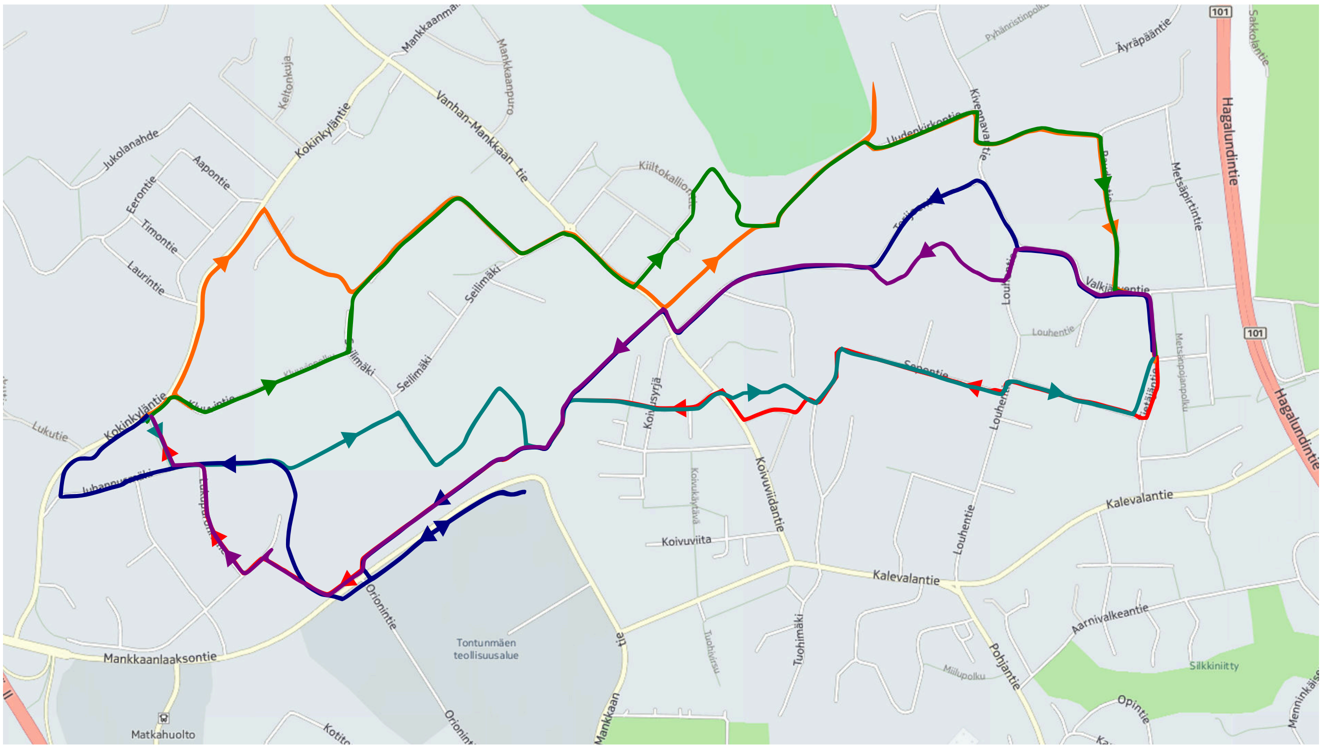


Fig. 28 The northern routes taken by the participants during beacon guidance in user study C

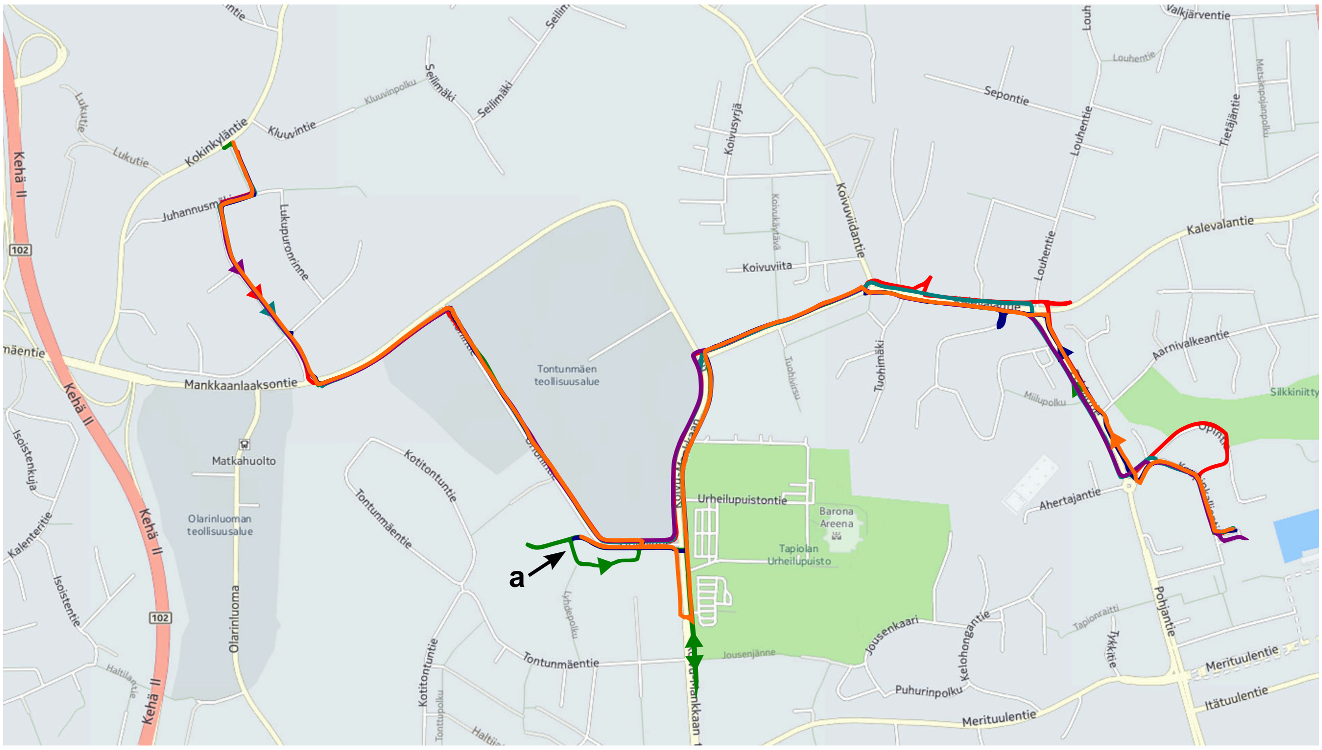


Fig. 29 The southern routes taken by the participants during route guidance in user study C. One problem with the implemented route guidance is illustrated at point *a*. There, participants arriving from the east naturally ended up on the cycle track on the south side of the road, instead of the north side. However, the track on the south side did not continue along the road that the participants were supposed to follow, but took another direction. After some confusion, the participants found there way back to the route that they were supposed to take

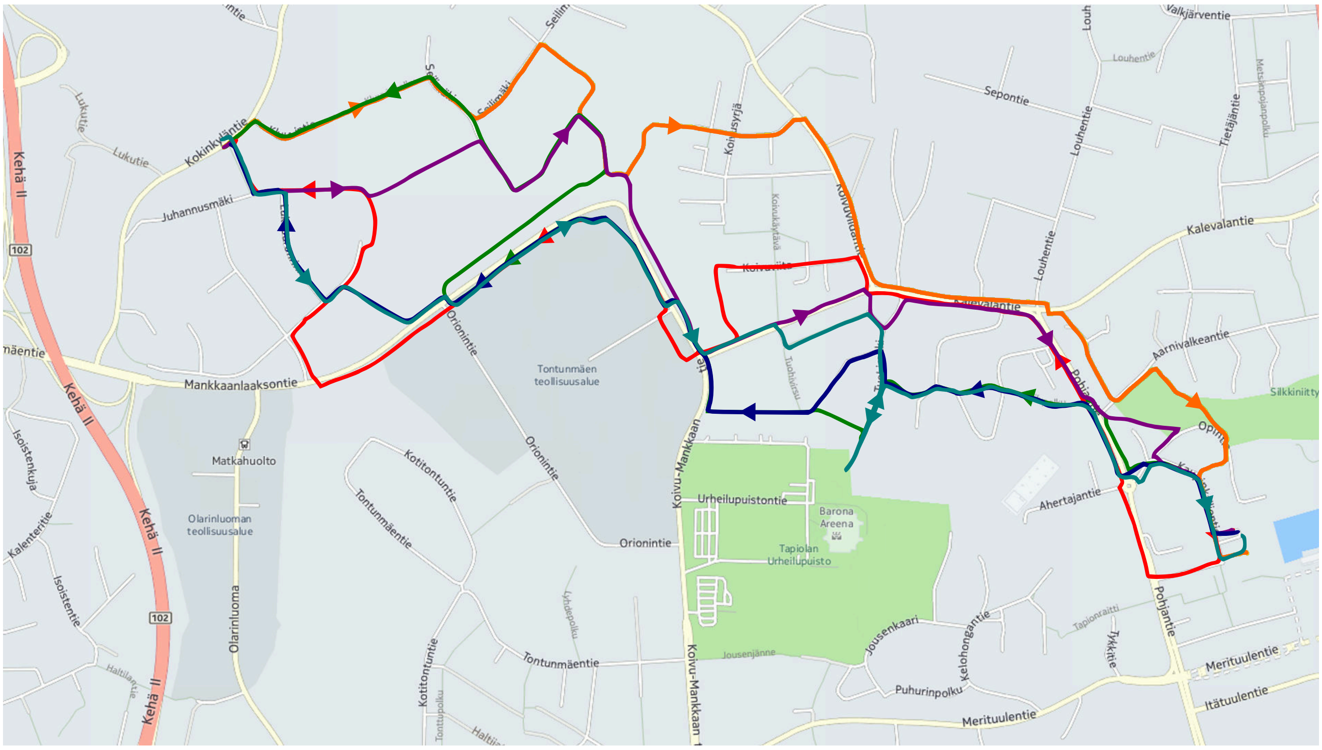


Fig. 30 The southern routes taken by the participants during beacon guidance in user study C