Research Article

Testing Two Tools for Multimodal Navigation

Mats Liljedahl,[1] Stefan Lindberg,[1] Katarina Delsing,[1] Mikko Polojärvi,[2] Timo Saloranta,[2] and Ismo Alakärppä[3]

[1] *Sonic Studio, The Interactive Institute, Acusticum 4, 941 28 Piteå, Sweden*
[2] *University of Oulu, PL 8000, Oulun Yliopisto, 90014 Oulu, Finland*
[3] *University of Lapland, P.O. Box 122, 96101 Rovaniemi, Finland*

Correspondence should be addressed to Mats Liljedahl, mats.liljedahl@tii.se

The latest smartphones with GPS, electronic compasses, directional audio, touch screens, and so forth, hold a potential for location-based services that are easier to use and that let users focus on their activities and the environment around them. Rather than interpreting maps, users can search for information by pointing in a direction and database queries can be created from GPS location and compass data. Users can also get guidance to locations through point and sweep gestures, spatial sound, and simple graphics. This paper describes two studies testing two applications with multimodal user interfaces for navigation and information retrieval. The applications allow users to search for information and get navigation support using combinations of point and sweep gestures, nonspeech audio, graphics, and text. Tests show that users appreciated both applications for their ease of use and for allowing users to interact directly with the surrounding environment.

1. Introduction

Visual maps have a number of advantages as a tool for navigation, for example, overview and high information density. Over the last years, new technologies have radically broadened how and in what contexts visual maps can be used and displayed. This development has spawned a plethora of new tools for navigation. Many of these are based on graphics, are meant for the eye, and use traditional map metaphors. The Google Maps application included in, for example, iPhone and Android smartphones is one example. However, visual maps are usually abstract representations of the physical world and must be interpreted in order to be of use. Interpreting a map and relating it to the current surroundings is a relatively demanding task [1]. Moreover, maps often require the users' full visual attention, disrupt other activities, and may weaken the users' perception of the surroundings. All in all, maps are in many ways demanding tools for navigation.

One major challenge for developers of navigation services based on smartphones is handling the inaccuracy of sensor data, especially GPS location data. The location provided by GPS in urban environments is often very inaccurate from pedestrians' perspective. Furthermore, the accuracy is heavily influenced by nearby buildings as well as other factors such as the positions of the satellites and weather.

This paper addresses the problems described above. The problem with current navigation tools' demands on users' attentional and cognitive resources was addressed using multimodal user interfaces built on a mix of audio, pointing gestures, graphics, and text. The aim was to study to what extent such interfaces could reduce the demands put on the users compared to more traditional navigation tools. To test the idea, two multimodal interfaces were developed. The main inputs to both applications are the device's GPS location and the direction in which the user is pointing the device. Both interfaces generate sound to indicate directions to targets and also present simple graphics and small amounts of text. The interfaces are built on the users' natural ability to locate sound sources and to follow a pointing arrow on the device screen and can to large degrees be used eyes-free.

This study was inspired and guided by four concepts and aims: minimal attention user interfaces, eyes-free interaction,

decreased cognitive loads on the users, and aesthetics of interaction. The study is also inspired by and based on a number of previous research efforts from several disciplines, including computer games, electronic navigation, and ubiquitous computing.

2. Background

Modes like auditive or haptic senses have been used for navigation applications in many studies. Examples include Tsukada and Yasymua [2], Frey [3], Amemiya et al. [4], Spath et al. [5], Loomis et al. [6], Kramer et al. [7], and Evett et al. [8]. But, as is often the case, the visual modality has drawn most attention when researching new interfaces for navigation. Also, as pointed out by McGookin et al. [9], work done on auditory navigation has primarily been geared towards people with visual impairments. There have been, though, a number of efforts developing auditory systems for navigation for sighted. *AudioGPS* by Holland et al. [10] is early work with spatial, nonspeech audio to convey information about the direction and distance to a target. *GpsTunes* by Strachan et al. [11] and *Ontrack* [12] by Jones et al. used spatially modified music to convey the same information. *Ontrack* plays music to lead the user towards a target destination; the music's spatial balance and volume indicate the directions the user should choose. A majority of test subjects were able to successfully navigate both a virtual and the physical world using the nonspeech audio provided by the system. *Ontrack* links to our own previous work on audio-based navigation in a virtual environment. *Beowulf* [13] showed that a soundscape together with a low-resolution graphic map is enough to present an entertaining and suitably challenging computer game. In *Audio Bubbles*, McGookin et al. [9] used audio to inform tourists about nearby points of interest. The users of the system can attend to or ignore the audio information. The aim of the Audio Bubbles is to promote a serendipitous or "stumble upon" type of navigation that is more targeted to exploration and experience than efficiency. The bearing-based navigation used in this study holds the potential to work in a similar way.

HaptiMap has produced a number of results related to the design, implementation, and evaluation of maps and location services that are more accessible through the use of several senses such as touch, hearing, and vision. See, for example, [14, 15]. Suitable angle sizes for pointing gestures were studied in [16]. *SoundCrumbs* [17] uses an interesting navigation method where a trail of virtual "crumbs" is laid out and the application helps a user to follow this trail via vibrotactile cues. The method can be described as based on bearings to a sequence of relatively close targets. The *PointNav* [18] prototype allows a user to both scan for points of interest (POIs) and to get guidance to selected POIs using a combination of pointing gestures, vibrotactile cues, and speech.

The works referred to above have all been successful in using multimodal interfaces to guide users to selected locations. The study described in this article continues this
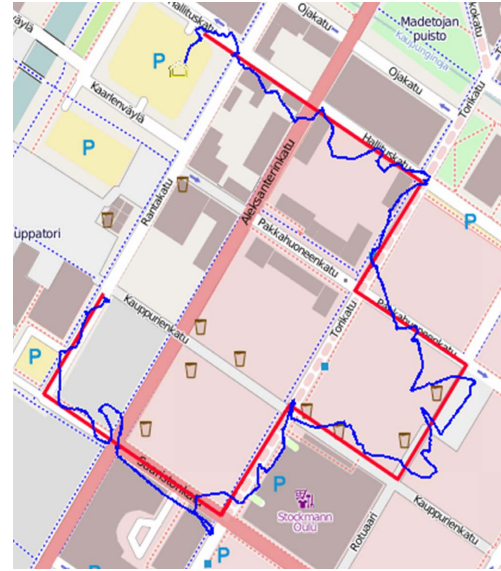

FIGURE 1: Route actually walked (red) and corresponding GPS locations logged (blue).

work and adds insights into the attentional and cognitive resources needed when using this approach on navigation.

Smartphone devices cannot supply as high accuracy in location and direction data as car navigators can. Decreased accuracy makes it troublesome to apply the turn-by-turn type of navigation used in car navigators to smartphone-based navigation applications for pedestrians. Another challenge is, as Pielot and Bol [19] point out, that pedestrians use navigation services in significantly different contexts compared to car drivers. Thus, it is important to find alternative navigation solutions for pedestrians. In a prestudy to the work reported here, one researcher walked a route in a city centre while logging GPS locations. When the logged data was compared to the route actually walked it was obvious that the logged locations often differed 30 metres or more from the actual locations. The map in Figure 1 shows the difference between the route actually walked (the thin red line) and the corresponding GPS locations logged (blue line).

Djajadiningrat et al. [20] argue that good interaction design should respect all of man's skills: cognitive, perceptual-motor, and emotional skills. This leads to interaction design where also what the user perceives with her senses and what she can do with her body become important in the design process. Hekkert [21] divides experience into three levels: aesthetic level, understanding level, and emotional level and sees aesthetics as "pleasure of the senses." It can generally be argued that aesthetics is a vital part of any user experience and is essential in developing useful, easy to use, and attractive products. The work reported here has strived to embody these ideas in the applications developed. The bearing-based guide function puts the users' cognitive and perceptual-motor skills at play in an attempt to overcome problems with fluctuating accuracy in GPS localization. But we also strongly believe that, at the same time, this promotes

a qualitatively different experience compared to turn-by-turn navigation along the lines of "serendipitous navigation."

3. Two Studies

Two studies were made. The first compared navigation using a paper map to navigation using a multimodal application. Two aspects of navigation were compared: the user's ability to follow a route and her awareness of the surroundings while navigating. The second study looked at users' reactions to using a multimodal application to find and to navigate to locations in a city environment. Two prototypic mobile applications were developed as tools for the studies. Both applications had multimodal user interfaces built on point and sweep gestures, spatial and nonspatial sounds, and text and simplistic graphics.

3.1. The First Study: The Audio Guide. The first study focused on providing answers to the following research questions and testing the corresponding hypotheses.

Q1: Do the users show and experience any difference in awareness and mental presence in the surroundings when using the multimodal application compared to using a map?

H1: Users will be more aware of and mentally present in the surroundings when using the multimodal application compared to using the map.

Q2: Do the test subjects perceive any difference in how mentally and physically demanding a navigation task is using the multimodal application compared to using a traditional map?

H2: The users will experience the multimodal application as less demanding mentally and physically compared to a traditional map.

The users' task was to navigate a predefined route by foot while at the same time looking for small signs with letters along the route. One-half of the route was navigated using a map and the other half using a multimodal application; the Audio Guide. The users were told to write down the letters in the order they found them along the route. Each route had seven or eight size A5 signs, each with a single black lowercase letter on white background. The letters did not form any intelligible word. Each sign was placed in a clearly visible location within 1–10 meters from the road. Each route was roughly 2 km long along roads, sidewalks, or bikeways and featuring 8-9 straight turns (Figure 2).

After an introduction based on a PowerPoint presentation the users were randomly given one of the navigation tools and asked to navigate the route alone. The multimodal application rendered sounds on top of environmental sounds, via loose-fitting headphones. Halfway through the route, the test leader met the users and gave them a questionnaire related to the navigation tool used. The users filled out the questionnaire and were asked to navigate the rest of the route using the other navigational tool. The questionnaire was based on the NASA Task Load Index (TLX) [22]. Since

the test at hand did not put any timely constraints on the users, the question about temporal demands in the original NASA TLX was replaced by the question "How attentive were you while performing the task?". The question about frustration in the original NASA TLX was rephrased "How irritated were you while you performed the task?".

To complement the NASA TLX, the test users also rated three statements on six-level Lickert scales from "Do not agree at all" to "Completely agree." Each statement concerning the multimodal application (Application) had a counterpart for the map (Map). The statements were as follows.

(1) *Application:* The Audio Guide was a good aid to find the way even if I did not look at it all the time, but, for example, kept it in my pocket.
    Map: The map was a good aid to find the way even if I did not look at it all the time, but, for example, kept it in my pocket.

(2) *Application:* To search for the correct way using the application's sound pointer was a powerful tool for navigation.
    Map: To search for the correct way by looking at the map and orienting it to the surroundings was a powerful tool for navigation.

(3) *Application:* I found it difficult to understand and use the Audio Guide.
    Map: I found it difficult to understand and use the map.

The test was conducted in the same way in three cities: in Oulu and Rovaniemi in Finland and in Piteå in Sweden. A total of 28 test users were recruited to the study. The test users' average age was 30 years, youngest 20, oldest 42, and median 27 years old. 14 were male and 14 female. Test users were students and staff at the universities in the three cities and volunteered to the test. The tests were performed in October and November 2010. Weather and daylight conditions differed between the tests. The tests in Oulu and Piteå were done in good weather conditions, in daylight, with no precipitation, and temperature above freezing. In Rovaniemi, half of the participants did the test in daylight and no precipitation and half in dusk and snowfall, temperature somewhat below zero.

3.1.1. Multimodal Application Used in the First Study. The Audio Guide application guided users along predefined routes using turn-by-turn navigation. The application had two distinct modes, follow and seek. Both modes primarily gave the user information via directional, nonverbal audio in stereo headphones.

In follow mode, users walked along the route waiting for instructions to turn left or right. The application tracked the user's location using the device's GPS system. The data supplied by the operating system was used without filtering or other manipulation. When the user was less than 30 meters from a waypoint, the application first played a notification sound alerting the user about the next instruction. When the user got closer than 20 meters from the waypoint, the

(a)

(b)

FIGURE 2: (a): Map showing one of the routes. (b): Placement of one of the letter signs.

application played an action sound indicating that the user should change course by turning left or right. All instructions provided by the application were short (0.3–3.2 s), nonverbal sound effects. The turn left and right signals were panned towards left and right ear, respectively.

Seek mode allowed users to detect the next waypoint from a distance. The application calculated direction and distance to the next waypoint using data from the mobile phone's GPS and compass sensors. Therefore, in seek mode the orientation of the phone affected the instructions. The direction to the waypoint was conveyed to the user through a graphical arrow on the device screen and a sound in the user's headphones. The left-right panning of the sound was continuously updated to point towards the waypoint and the user experienced the sound as coming from the waypoint. When the user pointed the device to the right of the waypoint, the sound was stronger in the user's left ear and vice versa. The distance to the waypoint was shown as text on the device screen.

The follow mode was intended as the default mode to be used without holding the mobile device in the hand and possible to use even when riding a bicycle. The seek mode was intended to be used if the user became uncertain about the direction to the next waypoint. The user was free to switch between the two modes at any time.

The application used in the test was implemented as a Java MIDP 2.0 [23] application and tested on Nokia E55 and 6210 Navigator running Symbian S60 3rd edition.

3.2. Results from the First Study

3.2.1. Quantitative Results from the First Study. The test did not proceed completely without problems. Of the 28 test users, a total of six did not complete the test according to the given instructions. In this study our focus was not on detecting or correcting the misuse of the navigation tools, therefore the results from these six test users are not included in the statistical analysis presented next. However, these cases are discussed in more detail later.

Table 1 shows the average percentage of letter signs found in correct order when using the multimodal application (App) and the map (Map), respectively. P values are calculated using the Mann-Whitney U test. Overall, the test subjects were able to find most of the letter signs along the routes.

TABLE 1: Results for finding letter signs.

Found signs	App (%)	Map (%)	Stddev app	Stddev map	P (2-tailed)
Average all cities	89	84	11.2	14.8	0.2983
Piteå	93	79	9.3	13.8	0.0316
Oulu	81	90	15.3	9.4	0.4237
Rovaniemi	90	85	5.1	20.0	0.9362

The results from the three cities vary. This can be explained by several factors. In Piteå, all users managed to follow their route and weather and daylight conditions were good. The result from this test shows a statistically significant difference ($P = 0.03$) in the percentage of signs found when navigating using the application (93%) compared to using the map (79%). The same difference is not significant in the results from the tests in Oulu and Rovaniemi.

Table 2 shows the NASA TLX results from the study. NASA TLX values can be interpreted in a straightforward fashion, for example, "1" indicates the smallest "Mental demand." The results indicate that the users did not consider the task especially demanding either mentally or physically, regardless whether they used the map or the application. In the results from the tests in Oulu and Rovaniemi, the users did not consider either of the navigational tools requiring much effort. There was significant difference in the results from Piteå, where the users considered using the map demanding noticeably more effort (mean 2.5) than using the application (mean 1.7). This is a statistically significant difference on a 5% level ($P = 0.05$). For the total result in effort needed, the difference between using the application and the map shows a P value of 0.1. Significance was calculated using the Mann-Whitney test. Given the cases where the user completed the task according to the instructions, neither of the navigation tools was reported as significantly irritating. Finally, there was no significant difference in the reported awareness of the surrounding world.

When comparing results from the three extra statements concerning the application to the corresponding results concerning the map, statement 2 shows significant difference. For this statement for the application, the mean value = 4.9 and standard deviation (stddev) = 0.95. For the corresponding statement for the map, the mean value equals

Table 2: NASA TLX results from study 1. Means from answers on scale 1–6.

Parameter	Application				Map			
	Total	Piteå	Oulu	Rov	Total	Piteå	Oulu	Rov
Mental demand	2.2	2.3	1.8	1.8	2.2	2.6	1.7	2.2
Physical demand	1.4	1.4	1.2	1.7	1.6	1.5	1.3	2.0
Performance	5.1	5.2	5.0	5.0	4.9	4.8	4.8	5.2
Effort	1.4	1.7	1.2	1.7	2.0	2.5	1.3	1.8
Irritation	1.4	1.3	1.7	1.2	1.6	1.8	1.2	1.8
Attentiveness	4.7	4.6	4.5	5.2	4.6	4.2	5.0	4.8

4.0 and stddev = 1.4. The Mann-Whitney U test revealed that the difference was significant ($U = 161$, $z = 2.26$, $P = 0.02$, two-tailed).

3.2.2. Qualitative Results from the First Study. In order to find qualities and dimensions of the multimodal application missed or overlooked by the research team, oral feedback was collected from the test subjects on all three test sites. Strengths and weaknesses of the concept were discussed and the test subjects were asked to convey their experiences for the trial period.

Over all, the applications user interface was perceived as intuitive and easy to understand and the application's sound design was generally appreciated. Some users perceived the action sounds, panning from the centre to the left and right ear, respectively, as less clear than if the panning was omitted and the sound played in only one ear. Several comments related to poor integration between sound and graphics in the application—what is heard should also have a graphical counterpart.

It was stated that the map attracted and captivated the users' eyes more than necessary for the navigation task at hand. The application on the other hand was said not to demand the users full attention more than just before turns at waypoints. When using the map, some users reported frequent feelings of uncertainty about their current position and the correct route. As a contrast, the users reported that they felt great confidence in the application showing the way in a trustworthy manner.

Due to varying accuracy in GPS positioning, action sounds were reported to play very early or very late at some waypoints, causing confusion. A weakness in the design was said to be that the users did not get any confirmation that they turned in the right direction at waypoints. Another related weakness was said to be that the application just played the action sound once.

It was commented that maps give overview but the application does not. Being able to provide the users with, for example, a sense of distance left to the next waypoint would be a useful enhancement. Three participants stated that given the choice between the application and a traditional map when in a foreign city, they would choose the application.

3.3. The Second Study: PING! Leveraging from the experiences acquired from the first study, the second study focused on test users' experiences of and attitudes towards two aspects of pedestrian city navigation. The first aspect studied was the use of a multimodal search function for finding information about points of interest. The second aspect was to study if a multimodal bearing-based guide function could help overcoming problems with varying GPS location accuracy that make turn-by-turn navigation troublesome for smartphone users. The research questions and corresponding hypotheses for the second study were as follows.

Q3: Can an interface based on a combination of point and sweep gestures, audio feedback and text be used to effectively find information about nearby points of interest?

H3: Users will be able to effectively find information about nearby points of interest using a search method based on a combination of point and sweep gestures, audio feedback, and text.

Q4: Can users effectively navigate to specified locations in a city using a guide function that is based on a combination of virtual, spatial sound sources, and a graphical arrow to indicate directions to targets and text to indicate distance?

H4: A majority of users can effectively navigate to specified locations in a city using a guide function that is based on a combination of virtual, spatial sound sources, and a graphical arrow to indicate directions to targets and text to indicate distance.

Q5: Can the interfaces described in Q3 and Q4 be effectively and successfully used despite varying accuracy in GPS location and compass directional data?

H5: Dividing the responsibility for finding the way to the target between the user and the application will help the user cope with varying accuracy in GPS positioning and electronic compass data.

The find and guide functions were implemented in a smartphone application designed for pedestrians. Both functions were based on point and sweep gestures for input and sound and simple graphics for output. The users' task was to use the find function to find directions to three target locations in the city. Then they should use the guide function to navigate there. In contrast to the first study, the approach was to let the users explore and choose the way to the target themselves, providing only information about distance and direction to targets.

The second study was conducted in August 2011 in Oulu with 24 users. The age range was 14 to 50 years, 15 users were female and 9 male. Each of the four test sessions lasted for three hours and followed the same structure. 22 of the test users were sent out two-by-two, forming 11 pairs. Two test users were sent out alone. Each pair of users had one smartphone and two headphones connected to the phone using a headphone output split adapter. Each group also had paper and pen to take notes on details from the target location to show that the correct location had been reached.

The structure of the test was the following. First 20 minutes introduction and instructions indoors followed by 10 minutes instructions and application demonstration outdoors. After this, the test users were sent out for 90 minutes for the actual test. After this the test users were brought back indoors where they filled out a questionnaire. The first part of this was the same version of the NASA Task Load Index used in the first study. One question was added: "How much did the application hinder your awareness of the world around you." The second part of the questionnaire had six statements relating to complementary aspects of the navigation experience. The users graded their answers on six grade Lickert scales ranging from "totally disagree" to "totally agree." In order to reveal aspects and qualities of the test not captured by the questionnaire all test subjects also took part in focus-group interviews.

3.3.1. Multimodal Application Used in the Second Study. The application used in the test had a database with points of interest (POI. Each database item held information about GPS location, street address, a photo, and some extra information about opening hours etc.)

The find function used point and sweep gestures as input and a combination of nonspeech sound, graphic icons, and text as output and was used in three steps. When using the find function, the user slowly swept the device back and forth (i.e. left to right and back). When the device was pointing towards a POI in the database the application played a short sound. The sound changed depending on the distance to the POI. The idea was to give users information about the density of POIs and distances to these in different directions. The next search step was called Fetch. The user pointed the device in some interesting direction and pressed the Fetch button. The application searched the database for POIs located in a sector with an angle −15 to +15 degrees from the device's current direction and up to a maximum distance of 2 km (Figure 3). Angles were selected based on the work by Magnusson et al. in [16].

POIs found by the Fetch function were presented in a scrollable list with name and distance from the user's current location. If the user wanted more detailed information about an item, she tapped the item in the list and the application switched to the item detail view. From this view, the user could get guidance from the current location to the item's location by tapping the "Guide me" button. In the guide view, the direction and distance to the target were presented. Based on this information the user herself decided the actual route to the target. The aim was to divide the responsibility

for finding the way to the target more equally between the application and the user compared to turn-by-turn navigation.

The guide function conveyed information about distance and direction to targets using directional audio in headphones, following the ideas of Robinson et al. [24]. As the user moved and turned, a sound moved in the users stereo image such that the sound seemed to come from the target. The more to the left of the target the user pointed the device, the more to the right ear the sound moved and vice versa. The sound was low-pass filtered as a function of distance to the target. Distance was also shown as a number on screen and direction was shown using a graphical arrow pointing towards the target.

The application was implemented on the Android platform and tests were done using Samsung Galaxy S2 devices.

3.4. Results from the Second Study

3.4.1. Quantitative Results from Second Study. All test users managed to find and walk to all the three target locations. Figure 4 shows the results from the NASA TLX. The graph does not show any high mental or physical loads on the users. Also, the users reported thinking they succeeded well in navigating the city using the application. They were not overly irritated while doing the task and the application did not hinder their awareness of the surroundings to any great extent while using it.

There are some differences in the results from the teenagers compared to the results from the adults. The teenagers reported succeeding somewhat better than the adults. A Mann-Whitney U test shows that the difference is statistically significant ($P = 0.02$). There are no significant differences for the other measures.

Figure 5 shows results from the six complementary statements in the questionnaire as the percentage of users in strong agreement with the statements (1 or 2 on the Lickert scales), in strong disagreement with the statements (5 or 6 on the Lickert scales), or showing a weak opinion (3 or 4 on the Lickert scales).

3.4.2. Second Study Focus-Group Interviews. All participants took part in focus-group interviews. The results are summarized below.

General Usefulness. Overall, the application was perceived as easy to use and to find the targets using the application's find and guide functions. It was noted that having information in sound leaves the eyes free for exploration. The guide function, showing only the direction towards the target, leaves the users free to choose their own way. This, in turn, was said to have benefits.

Comparison to Maps and Car Navigators. When using maps, users stated that they are often unsure if they interpret the map correctly in relation to the environment. Several users appreciated the application for its ability to know where you are and to show you the direction in the

FIGURE 3: The Fetch function retrieved information about POIs in a 30-degree sector in front of the device.
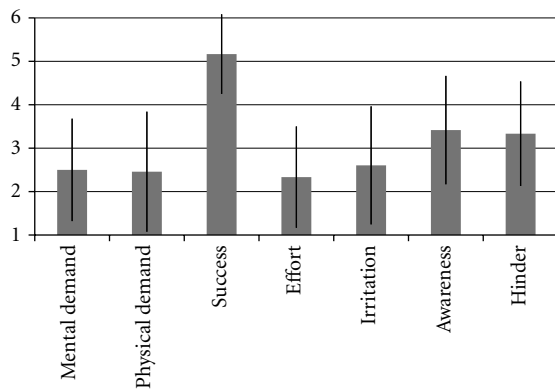

FIGURE 4: NASA Task Load Index. Average of all users.

physical environment. Several of the teenagers referred to car navigators as annoying.

Sound Feedback. The audio feedback from the application received mixed opinions. Some users reported they did not actively listen to the sound feedback at all, instead using the onscreen graphic and textual information to search for and navigate to targets. Other users appreciated the ability to use the application "eyes-free," just listening to the audio feedback. The guide function's spatial audio and the ability to determine the number of POIs in some given direction by sweeping were found useful by these users. The ability to visually observe the surroundings while using the application was a good feature mentioned by several users.

Some users asked for greater diversity between the different sounds in order to more easily discriminate between them and their different meanings. At some occasions, the sound from the application was drowned by background noise from traffic or machines. To some users the application did not convey enough information through audio about direction (left/right) or distance to target. Using speech to give the information "turn left" and "turn right" was suggested as a solution.

Balance between Sound, Graphics, and Text. Overall, the users reported having relied to the graphical and textual information on the screen more frequently than to the

information conveyed by the sound feedback. The sound feedback was useful to get information about direction to target, but in order to get an idea about distance to the target, the users still had to rely primarily on the onscreen text information. The users also reported relying on the onscreen information when the sound from the application was drowned by background noise.

4. Discussion

4.1. The First Study. The tests could not verify the hypothesis that a multimodal application would let the users be more aware of and mentally present in the surroundings compared to a traditional map (H1). The results from the three test sites differ somewhat. The tests in Sweden were performed with the least disturbances from weather and other conditions. The results from these favourable conditions show a significant difference in number of letter signs found using the multimodal application compared to using the map. Tested under less favourable conditions this difference is not significant. The difference in overview the map gave compared to what the application gave probably affected the results. Some users were for example so familiar with the surroundings that a short glance at the map to check where to turn next was enough for walking several hundred meters without worrying about getting lost.

The tests could not fully verify the hypotheses that a multimodal application would put a lower mental and physical load on the users compared to a traditional map (H2). The NASA Task Load Index does not show any significant differences in the mental or physical demands the two navigation tools put on the users. However, users reported that the overall effort needed to perform the task was significantly lower using the application compared to using the map.

It can be concluded that navigation based on primarily turn-by-turn-based instructions similar to car navigators is not sufficient in pedestrian settings. The primary cause for this is the inaccuracy of the GPS sensor, even 30 meter errors are common. Often this leads to situations where turning signals come too early or too late in relation to the turning point. This is a significant problem. If the signal comes too soon, the user may end up choosing the wrong turn. If too late, the user may get confused over whether she should
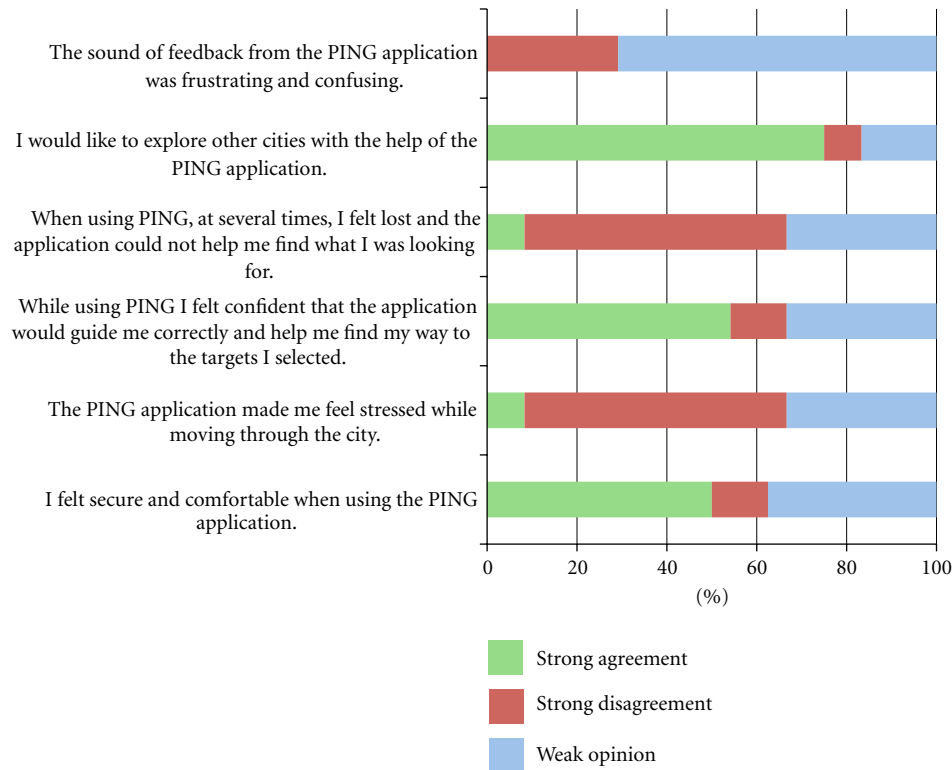
FIGURE 5: Users level of agreements to complementary statements in the questionnaire.

return to the previous intersection or continue to the next. There are also many intersections where simple 90-degree turn signals are not sufficient to signify which direction the users should go. Another conclusion is that during the first minutes of using a new application, learning the application might absorb some users. This might in turn pose security risks to these users.

4.2. The Second Study. The results from the second study indicate that the hypothesis holds true that the users are able to find information about nearby points of interest using a search method based on point and sweep gestures, audio feedback, and text information (H3). The results also indicate that the hypothesis holds true that a majority of users will be able to find routes and navigate to selected points of interest with the help of a guide function showing (only) the direction and distance to the selected POI using a combination of spatial audio, a graphic, onscreen pointer, and text (H4).

The turn-by-turn navigation used in the first study imposed high demands on GPS-location accuracy and required the users to stick to major roads. The target bearing-based navigation method used in the second study was designed to overcome this problem by dividing the responsibility for finding the way more equally between the user and the device. Despite large documented variances in GPS position accuracy, all users managed to successfully use the target-bearing guide to navigate to the targets. This suggests that the hypothesis holds true that dividing the responsibility for finding the way to the target between the user and the application helps the user cope with varying accuracy in GPS positioning and electronic compass data (H5).

The results from the NASA Task Load Index suggest that using the application did not put heavy mental or physical loads on the users. The users reported that they thought they succeeded well in navigating the city using the application and they were not overly irritated while doing so. These results indicate that the application did not hinder the users' awareness of the surroundings to any greater extent. The same is true for how aware of the surrounding world the users were while doing the task. Informal statements from the test users indicate that for many of them, the test situation was their first encounter with bearing-based navigation, for some also with smartphones and touch screens. This suggests that learning curve effects are parameters affecting the results. Being occupied with learning a new user interface and interpreting auditory and graphic feedback from it may have influenced the awareness of the surroundings and how much the test users perceived the application to hinder this awareness. Together the results give further support for hypotheses H3, H4, and H5.

4.3. Overall Discussion. Two prototype applications for navigation and information retrieval were developed. These applications did not rely on maps presented on graphical user interfaces but were instead based on users' innate abilities to use point and sweep gestures to indicate directions

and to use directional hearing to locate sound sources. The graphical user interface presented general information about points of interests, distances to them, and directions to selected POIs. Point and sweep gestures made by the user and spatial audio were in the main roles in these user interfaces.

The applications were evaluated in two studies. Based on the results from these it can be argued that applications for navigation featuring-multimodal user interfaces hold the potential to help users find the way, while at the same time leaving them more free to experience and explore the surrounding environment compared to navigation using traditional maps. The studies show several similarities. The users perceived themselves as successful in fulfilling the navigation tasks and the applications did not put high mental or physical loads on the users. Both quantitative and qualitative data indicate that both applications tested were generally appreciated for their ease of use and overall efficiency.

The application tested in the first study used turn-by-turn navigation. In the study it was found that this type of navigation is sensitive to variances in GPS location and compass-direction accuracy. The target-bearing based navigation used in the second study divides the responsibility for the navigation task more equally between the user and the device. The user is free to explore the city; the application helps the user to stay on a route towards the target, without distracting the exploration too much. The second study revealed that this type of navigation is less sensitive to varying accuracy in GPS location and compass-direction data. That is, the bearing-based application is more robust than the turn-by-turn application, because there are no waypoints requiring accurate calculations from the application and swift reactions from the user. Moreover, single errors have only a momentary effect as calculations are always based on the latest sensor values. A user does not need to react to instructions exactly at waypoints, but to check periodically the direction and distance to the target and decide the route in the environment herself based on this information. Finally, since there is no predefined, correct route there is no need to check whether the user is on that route and correct her if not.

For some users, sound feedback can be hard to interpret and to make use of. The application used in the second study did some integration between sound and graphics/text, so when a sound was played some onscreen information changed at the same time. Integrating several modalities might produce good user experiences for larger user groups and facilitate learning the application.

Several of the adult users expressed that they found the turn-by-turn navigation used in car navigators better and more efficient than the target bearing-based navigation used in the second study. This preference towards turn-by-turn navigation might simply be because of their greater and more long-time experience with this type of devices. Despite the inherent technical problems in implementing turn-by-turn navigation in pedestrian settings with GPS technology, turn-by-turn navigation may still prove a useful mix with other navigation methods. This observation is supported by the first study, where the majority of the testers were still

able to navigate the given route with primarily turn-by-turn audio instructions. Moreover, some waypoints might be necessary when a large obstacle like a river or a railway station requires a detour. On the other hand, guidance based on target bearing was successful and all test users navigated successfully to all given target locations. This guidance style can be seen as a very promising approach to cope with GPS inaccuracies.

The implemented applications also serve as examples how good user experiences do not require the state of the art technology like displays with the best resolutions and largest sizes, but instead careful design of multimodal, natural user interfaces can provide good results.

Acknowledgments

References

[1] A. M. MacEachren, *How Maps Work: Representation, Visualization, and Design*, The Guilford Press, 1995.

[2] K. Tsukada and M. Yasymua, "ActiveBelt: belt-type wearable tactile display for directional navigation," in *Proceedings of the 6th International Conference on Ubiquitous Computing (Ubicomp '04)*, pp. 384–399, Springer, 2004.

[3] M. Frey, "CabBoots: Shoes with integrated guidance system," in *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, pp. 245–246, ACM, February 2007.

[4] T. Amemiya, H. Ando, and T. Maeda, "Lead-me interface for a pulling sensation from hand-held devices," *Transactions on Applied Perception*, vol. 5, no. 3, article 15, pp. 1–17, 2008.

[5] D. Spath, M. Peissner, L. Hagenmeyer, and B. Ringbauer, "New approaches to intuitive auditory user interfaces," in *Proceedings of the Conference on Human Interface: Part I (HCII '07)*, M. J. Smith and G. Salvendy, Eds., vol. 4557 of *Lecture Notes in Computer Science*, pp. 975–984, 2007.

[6] J. M. Loomis, J. R. Marston, R. G. Golledge, and R. L. Klatzky, "Personal guidance system for people with visual impairment: a comparison of spatial displays for route guidance," *Journal of Visual Impairment and Blindness*, vol. 99, no. 4, pp. 219–232, 2005.

[7] R. Kramer, M. Modsching, and K. Ten Hagen, "Development and evaluation of a context-driven, mobile tourist guide," *International Journal of Pervasive Computing and Communications*, vol. 3, no. 4, pp. 378–399, 2007.

[8] L. Evett, S. Battersby, A. Ridley, and D. Brown, "An interface to virtual environments for people who are blind using Wii technology—mental models and navigation," *Journal of Assistive Technologies*, vol. 3, no. 2, pp. 26–34, 2009.

[9] D. McGookin, S. Brewster, and P. Priego, "Audio bubbles: employing non-speech audio to support tourist wayfinding,"

in *Proceedings of the 4th International Conference on Haptic and Audio Interaction Design (HAID '09)*, pp. 41–50, Springer, 2009.

[10] S. Holland, D. R. Morse, and H. Gedenryd, "AudioGPS: spatial audio navigation with a minimal attention interface," *Personal Ubiquitous Computing*, vol. 6, no. 4, pp. 253–259, 2002.

[11] S. Strachan, P. Eslambolchilar, and R. Murray-Smith, "GpsTunes—controlling navigation via audio feedback," in *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '05)*, pp. 275–278, ACM, September 2005.

[12] M. Jones, S. Jones, G. Bradley, N. Warren, D. Bainbridge, and G. Holmes, "Ontrack: dynamically adapting music playback to support navigation," in *Personal and Ubiquitous Computing*, vol. 12, pp. 513–525, Springer, 2008.

[13] M. Liljedahl, N. Papworth, and S. Lindberg, "Beowulf: an audio mostly game," in *Proceedings of the 4th International Conference on Advances in Computer Entertainment Technology (ACE '07)*, pp. 200–203, June 2007.

[14] D. McGookin, C. Magnusson, M. Anastassova, W. Heuten, A. Rentería, and S. Boll, *Proceedings from Workshop on Multimodal Location Based Techniques for Extreme Navigation*, Helsinki, Finland, 2010.

[15] M. Anastassova, C. Magnusson, M. Pielot, G. Randall, and G. B. Claassen, "Using audio and haptics for delivering spatial information via mobile devices," in *Proceedings of the 12th International Conference on Human-Computer Interaction with Mobile Devices and Services (Mobile HCI'10)*, pp. 525–526, ACM, September 2010.

[16] C. Magnusson, K. Rassmus-Gröhn, and D. Szymczak, "Angle sizes for pointing gestures," in *Proceedings of the Workshop on Multimodal Location Based Techniques for Extreme Navigation*, Helsinki, Finland, 2010.

[17] C. Magnusson, B. Breidegard, and K. Rassmus GrÖhn, "Soundcrumbs—hansel and gretel in the 21st century," in *Proceedings of the 4th international workshop on Haptic and Audio Interaction Design (HAID '09)*, 2009.

[18] C. Magnusson, M. Molina, K. Rassmus-Gröhn, and D. Szymczak, "Pointing for non-visual orientation and navigation," in *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries (NordiCHI '10)*, pp. 735–738, ACM, October 2010.

[19] M. Pielot and S. Bol, ""In fifty meters turn left": why turn-by-turn Instructions fail pedestrians," in *Proceedings of the Workshop Using Audio and Haptics for Delivering Spatial Information via Mobile Devices (MobileCHI '10)*, Lisbon, Portugal, 2010.

[20] T. Djajadiningrat, S. Wensveen, J. Frens, and K. Overbeeke, "Tangible products: redressing the balance between appearance and action," in *Personal and Ubiquitous Computing 8*, pp. 294–309, Springer, London, UK, 2004.

[21] P. Hekkert, "Design aesthetics: principles of pleasure in design," *Psychology Science*, vol. 48, no. 2, pp. 157–172, 2006.

[22] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): results of empirical and theoretical research," in *Human Mental Workload*, pp. 139–183, 1988.

[23] Java MIDP 2.0, http://jcp.org/aboutJava/communityprocess/final/jsr118/index.html

[24] S. Robinson, M. Jones, P. Eslambolchilar, R. Murray-Smith, and M. Lindborg, ""I did it my way": moving away from the tyranny of turn-by-turn pedestrian navigation," in *Proceedings of the 12th International Conference on Human-Computer Interaction with Mobile Devices and Services (Mobile HCI '10)*, pp. 341–344, ACM, September 2010.

Composition Comments

1. We added the highlighted part and made the highlighted changes in the provided addresses. Please check, unify the language of the second address, and provide the name of division or department to the addresses lacking ones.

2. We made the highlighted change according to the list of references.

Author(s) Name(s)

It is very important to confirm the author(s) last and first names in order to be displayed correctly on our website as well as in the indexing databases:

Author 1
Last Name: Liljedahl
First Name: Mats

Author 2
Last Name: Lindberg
First Name: Stefan

Author 3
Last Name: Delsing
First Name: Katarina

Author 4
Last Name: Polojärvi
First Name: Mikko

Author 5
Last Name: Saloranta
First Name: Timo

Author 6
Last Name: Alakärppä
First Name: Ismo