Navigating by audio-based probing and fuzzy routing

Mikko Polojärvi
University of Oulu
Erkki Koiso-Kanttilankatu 3
FI-90014 University of Oulu
+358 294 482805
mikko.polojarvi@ee.oulu.fi

Timo Saloranta
University of Oulu
Erkki Koiso-Kanttilankatu 3
FI-90014 University of Oulu
+358 294 482805
timo.saloranta@ee.oulu.fi

Jukka Riekki
University of Oulu
Erkki Koiso-Kanttilankatu 3
FI-90014 University of Oulu
+358 294 482798
jukka.riekki@ee.oulu.fi

ABSTRACT

In this paper, we present a navigation application, PING!, with a unique combination of characteristics. First, a user can study her/his environment by probing, or sweeping the phone horizontally, upon which points of interest in the pointed directions are indicated with audio feedback. Second, PING! utilizes fuzzy routing. Turn-by-turn instructions are not given and maps are not shown. Instead, the direction of the target is indicated with a simple arrow and binaural audio and the user is free to decide the exact path to take based on this direction information and the environment surrounding her/him. The application, realized in Android devices, was tested in field tests that verified the feasibility of the application concept and revealed directions for further improvement.

Categories and Subject Descriptors

J.0 [**Computer Applications**]: General – *mobile navigation, audio interface.*

General Terms

Design, Experimentation, Human Factors.

Keywords

Navigation, eyes-free interface, audio interface, fuzzy routing, pointing, non-visual.

1. INTRODUCTION

Mobile phones and other mobile devices are nowadays used throughout the day, from dawn to dusk, for various tasks – making calls, managing schedules, browsing the Internet, playing games, and so on. As more and more of these devices are equipped with positioning functionality, location-based services are becoming increasingly popular. In fact, phones are replacing paper maps and dedicated GPS navigation devices. However, navigation is still based on exact, turn-by-turn directions that are shown on street maps or spoken aloud by the application.

In this paper, we describe a novel navigation application, PING!, that discards visual street map view and turn-by-turn signals. Maps are discarded in both target selection and navigation. A user does not study points of interest (POIs) from a map but by *probing* the local environment. In probing, a user sweeps the phone horizontally (i.e. left to right and back). Audio feedback informs at each pointed direction the number of and distances to the selected types of POIs. The user can request more information about the discovered POIs to the phone screen and select a target based on that information.

Once a target has been selected, PING! guides the user to the target by *fuzzy routing*: the application presents only direction and distance to the target and leaves the selection of the exact path to the user – exact turn-by-turn instructions following a pre-calculated route are not given. Moreover, eyes-free, audio-based, binaural guidance is emphasized and GUI is used merely to present the direction of the target with an arrow and the geodesic distance to the target – maps or verbal instructions are not shown. Both probing and fuzzy routing use GPS for positioning and electronic compass to determine the direction the phone is pointed to. Stereo headphones provide audio feedback.

Our motivation for developing this application is to decrease the amount of attention the application requires. We hypothesize that when a user is relieved from the task of interpreting a map and textual instructions, she/he can observe and experience her/his surroundings more freely – have her/his focus of attention on the environment and be more aware of and present in the environment. Moreover, as pedestrians and cyclists often do not have as strict constraints on path selection as drivers have, fuzzy routing gives them more freedom in selecting the path. Finally, we believe probing to produce more serendipitous navigation: users do not stick with the places they know beforehand, but learn new ones by probing. We discuss later in this paper how well the application implemented on Android phones met these expectations in the field tests.

Others have used non-speech audio for guiding users towards virtual beacons placed in the environment [1] and for indicating the direction and distance to the next waypoint [2]. When compared to the related work, our main contributions are in probing, fuzzy routing, sensor usage, and how information is presented with binaural audio and sound properties. The audio design and the field tests are discussed also in [3], [8]. We do not present the details of the audio design in this paper but describe in more detail the operation principle and implementation of the PING! application, present and analyze new information about the field tests, and discuss how the application could be improved. Moreover, we report our experiences on using GPS, magnetometer, and accelerometer together in order to provide real-time feedback to the handling of an Android phone.

The rest of the paper is organized as follows. First, related work is presented in the second section. Third section presents the operation principle of the realized application and fourth its implementation, including related challenges. Fifth section describes the field tests, followed by discussion in the sixth section. The paper is concluded in the seventh section.

AcademicMindTrek'13, October 01-04 2013, Tampere, Finland.

2. RELATED WORK

As such, the idea of using language free audio signals to provide navigational aid is not new. Magnusson et al. [1] describe an application, Soundcrumbs, with which a user can form trails by placing virtual audio beacons emitting different sounds to the environment. Soundcrumbs uses only sound volume as a modality of conveying information regarding the correct direction the user should proceed to. While Soundcrumbs uses only mono sound and is meant to be used to form quite precise paths, it does demonstrate that language free audio alone can be a viable mean of navigation information. This supports our findings.

AudioGPS [2] is designed to provide a Minimal Attention User Interface (MAUI) [10] for navigation with non-voice audio to guide the user. It uses panning of the sound to indicate the bearing to the next navigation waypoint and an audio pulse count and rapidity as an indication of distance to the waypoint. AudioGPS differs from PING!, as it calculates the bearing to the target from GPS track, i.e. the route the user's GPS device measured. Due to the way GPS works, this can be problematic, especially in city environments where buildings cause errors to GPS data. Moreover, probing cannot be realized with GPS only.

The gpsTunes prototype [11] combines an mp3 player implemented on a PocketPC with an external navigation module. This system guides the user to a pre-defined location using song volume panning to indicate bearing to the target and general volume level adjustment to indicate the distance to the target. The reported user behavior from user tests in [11] is similar to our observations from user tests with PING! application. In both cases, initial reaction of the users is to rotate around without moving to get an idea how the volume panning works and to get an idea of the general direction of the navigation target.

In close relation to [11], Jones et al. [12] studied the use of adjusting the channel balance and volume level to guide users to target destinations. The ONTRACK prototype uses music as sound which is then volume adjusted to enable eyes-free navigation. ONTRACK was tested both in laboratory conditions and also outdoors with recruited test users. The general observations from tests show that adjusting stereo audio source's channel volumes is an effective way of finding a specified navigation target. However, an issue concerning the usage of music as the audio source for navigation was noted. Jones et al. reported that normal volume fading at the end of the song confused some of their users because they thought the fading meant they were going in the wrong direction. The audio signals in PING! application do not have this problem.

The old GPS navigation devices that did not present high-resolution maps resemble PING!, as direction and distance were emphasized in these devices' user interfaces. Also special devices like dog tracking collars present similar arrows as PING!, though they track a single mobile target. Moreover, the probing sounds were inspired by radar user interfaces. Garmin's GPS-enabled running sports watch [13] is an example of a GPS-device with similar graphical guidance as PING! has: an arrow and distance reading to a preset target location.

A trend in commercial applications, mainly due to iPhone's lack of Wi-Fi positioning, is the emergence of navigation applications that do not use sensor data at all. For example, Fastmall (fastmall.com) and aisle411 (aisle411.com) provide turn-by-turn walking instructions, but they require their user to define both the starting point and the target. Whatamap (whatamap.com) is also an indoor navigation application which can operate without sensor data. In such a mode, it essentially acts as an interactive map without guidance features. All these applications are based on the assumption that the user follows exactly the given path. PING!, on the other hand, lets the user to decide the exact path and instructs the user on that path.

Blindsquare (blindsquare.com) is a new application for visually impaired that uses GPS and compass, as PING! does, and describes the environment at the pointed direction using speech synthesis. Blindsquare resembles the probing phase of PING! (i.e. sweeping), except that speech is used instead of sounds.

3. OPERATION PRINCIPLE

Traditional GPS devices and navigation applications provide turn-by-turn navigation with verbal directions. These applications require that the user searches the desired target through levels of menus or writes the target's name to a search field. Moreover, the applications require the users' focus of attention – users have to focus their gaze to the phone display to check the path.

As described in the introduction, PING! application approaches navigation from a different perspective. When starting the navigation, the user is presented with a screen listing available categories; each with a large symbol and a name. After at least one category has been selected, the user can start to probe the environment, that is, to sweep the phone horizontally, from left to right and back. While doing this, the user can hear the density and distances of POIs falling into the selected categories. The audio feedback is similar in nature to a Geiger counter, providing a beeping sound for each target found at the direction of the sweep. The distance to the target is divided into three categories: near, middle and far, each being represented with a different sound effect.

When the user has found a direction containing a satisfying number of targets, she/he can do the fetch action by pressing a "Fetch" button. The device responds by presenting a list of targets matching selected categories in the direction to which the user was pointing the device. For each target, a category symbol, name, and distance are presented. Target selection is based on *fetch sector*; a sector centered at the selected direction (Figure 1). In order to mitigate the problem of the direction-to-target information becoming increasingly inaccurate in close distances due to GPS inaccuracy, the fetch cone is projected from a certain distance behind the user, with targets positioned behind the user filtered out from the results.

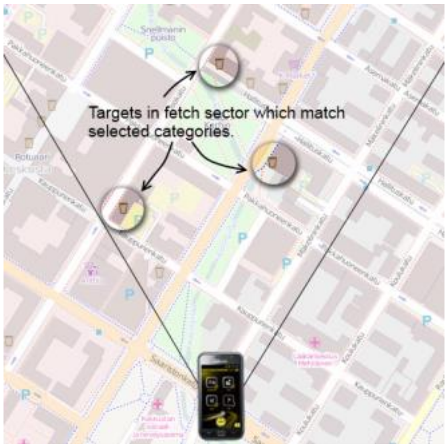

Figure 1. The fetch sector.

When the user sweeps the device in the selected fetch sector, the application provides both visual and auditory signals when the user is pointing the device directly towards a target in the list. The auditory signal is the same used in the previously explained category selection screen (Figure 2). The visual signal is realized by flashing the name of the POI in question on the GUI. The user can repeat the fetch action several times, always creating a fresh list of targets at the device's direction. Clicking a target in the list opens a detailed information screen about the target. For most targets, the detailed information consists of postal address, contact information, opening hours, photograph of the main entrance and a short description.

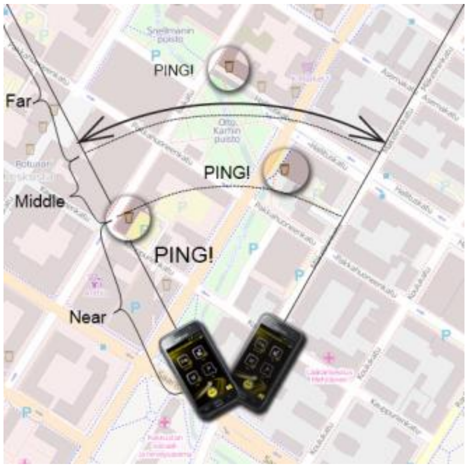

Figure 2. Sound feedback during a sweep motion.

Happy with her/his choice, the user can proceed to the navigation screen by pressing the "Guide me there" button. The navigation screen (Figure 3) shows the target's name, distance to the target (geodesic, "as the crow flies" distance), and direction to which the user needs to proceed in order to reach the target. A binaural guiding audio signal is repeated at set intervals (but kept mostly silent). While the user moves around in the city or sweeps the phone, besides keeping the direction arrow and distance updated, the stereo channel volumes of the phone are also adjusted so that the sound seems to be coming from the direction of the target (the direction the phone is pointed to is the "straight ahead" direction). When the user has reached the target, the audio signal "target reached" is played, and the guiding signal is stopped.

Figure 3. Navigation screen.

The device is capable of determining the compass reading even while the device is held vertically, for example in the user's front shirt pocket. This means that the user does not need to hold the device in her/his hand in order to receive the audio-based guidance toward the target. Moreover, to help the user to orient her/himself and detect magnetic disturbance in the compass, an arrow pointing north is presented on most screens. Finally, an icon signifying the current GPS connection state is shown in the upper right corner. This information is not communicated to the user via sound, but requires studying the device screen.

4. IMPLEMENTATION

4.1 Mobile Devices

The prototype was implemented for Android devices for Android OS API level 10 (version 2.3.3). The specific phone models were Samsung GT-i9000 (Galaxy S), Samsung GT-i9001 (Galaxy S Plus) and Google / Samsung GT-i9020 (Nexus S). The application utilizes magnetometer, accelerometer and GPS sensors and uses Internet connection to download POI data from a custom server.

4.2 Compass

Compass values are calculated from magnetometer and accelerometer readings with the assumption that the magnetic field corresponds to the Earth's magnetic field and the accelerometer reading to the direction of gravity. The calculated value is adjusted for declination, which depends on the user's location. The magnetometer values are adjusted to show the same compass direction when the device is held flat (display towards the sky), upright (display towards the user), and in any position between these two.

Besides the known problem of magnetic interference from certain objects (e.g. power cables, loudspeakers, displays and machines) affecting the magnetometer readings, another problem relates to the sweep feature necessitating swinging the phone. Swinging affects the accelerometer readings, which in turn affects which direction is calculated to point downward. This effectively results in greater computed sweeping motion than is actually performed, triggering targets than are not within the actually swept area. This may lead to reduced user experience if the user cannot trust the sweeping actions to match with reality. Utilizing gyroscope and modern sensor fusion techniques would mitigate these problems with the cost of higher power consumption, but because the used Galaxy S models did not feature a gyroscope, the problem was tackled by low-pass filtering the accelerometer readings [5]. However, this method eliminates the problem only partially, and works to increase the delay in reacting to changes in the phone orientation. On the other hand, PING! does not necessitate sweeping the phone especially fast; as long as the phone is swept sufficiently slowly, the problem is hardly noticeable.

4.3 GPS

PING! is an improved version of our earlier navigation tool discussed also in [3]. In short, the purpose of this earlier application was to study the feasibility of GPS and compass, as well as the suitability of speech-free audio as a modality of conveying navigation information. The application guides the user by playing via headphones sound signals that indicate when to turn left or right. However, the unpredictably inaccurate GPS data made it difficult to time the turning signals correctly. From time to time, the signals would come too early or too late. When the signal came too early and there were multiple intersections close to each other, the users reported being unsure which intersection to choose. When the signal came even a couple of meters after the intersection, the users reported being confused on whether they were supposed to backtrack their steps or not. In some cases where the road would end before hearing the signal, some users ended up making an extra turn. The users also had differing opinions on whether, for instance, a 45 degree turn constitutes a turn in the application.

A test walk shown in Figure 4 [3] demonstrates this problem. The recorded GPS track differs noticeably from the actual route walked, especially around buildings with three stories or more, i.e. in common city environment. This error is more pronounced when closing the distance towards Earth's poles, due to the GPS satellites getting located closer to the horizon.

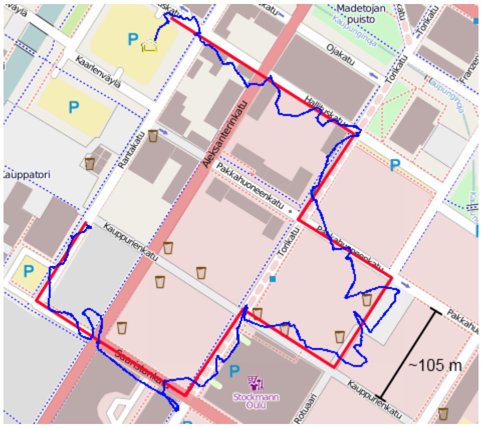

Figure 4. A typical GPS track. Red straight lines: the actual route. Blue zigzag pattern: the route recorded by the GPS.

With car navigators the usual practice is to match the location obtained from sensors to a map or a known network of roads [6], [7]. However, with bicyclists and especially pedestrians this approach can be considered significantly less feasible, since they are less restricted to traversing along known roads or paths.

The earlier application featured also a directional arrow on the mobile device screen, pointing toward the next waypoint along the route. The field tests revealed that the arrow points toward distant targets quite reliably, but gets increasingly inaccurate the closer the user got to the target. The GPS error started to dominate when the distance decreased under 30 meters, but on the other hand, then users were already able to locate the target visually [3]. These experiments led us to rethink our navigational approach and to abandon turn-by-turn instructions. Instead, we decided to focus mainly on informing the user of the correct direction when away from the target, and letting the user find the actual route her/himself. For the last meters, the application would just provide assistance in visual locating.

4.4 POI content

Upon starting the application, a database of POIs is downloaded from a remote server, both of which were designed for the field tests in mind. For each POI, the database stores information about the type of the establishment, its location, opening hours, short description, photo of the entrance and various related audio clips.

Another possibility would have been to acquire the data from an existing online service, for example, OpenStreetMap. A custom XML-based database was selected instead, because we required direct control over the contents of the database and did not want possible inaccuracy of the online search results to affect the test results. Another reason was our plan to develop characteristic sounds for each POI and use the server to distribute them to the mobile devices, although this was not tested yet in these field tests. Overall, even though the solution necessitated providing the contents for the database manually, it was found sufficient for the research task at hand.

4.5 Sound Playback

Sound feedback is played when a user probes the environment in the sweeping mode and during the actual navigation. During navigation, directional audio is played to indicate the distance and direction to the current target. As the Android API level 10 lacked proper support for 3D audio, directional audio was implemented by simply adjusting stereo sound volumes as a function of angle between the mobile device orientation and the direction to the target (Figure 5). The dashed lines in the figure visualize an example situation; indicating the directions of north and the target. The resulting volumes are illustrated in Figure 6, positive angle indicates that the target is to the right of the user.

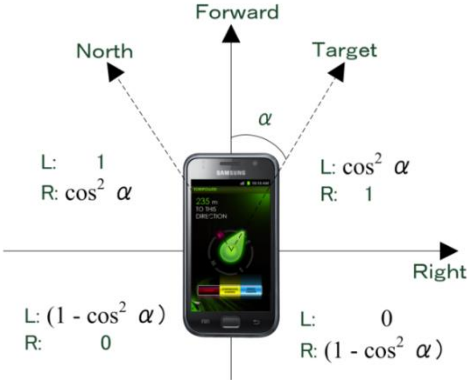

Figure 5. Sound channel volumes as a function of α, the angle between the user's orientation and the direction to the target.

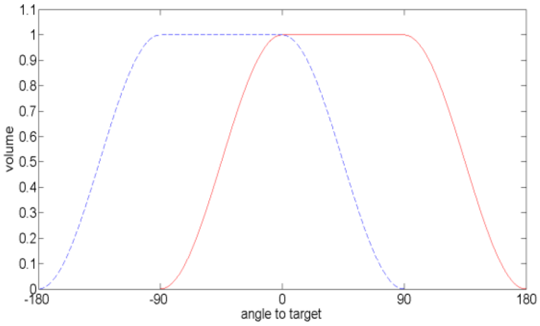

Figure 6. Sound channel volumes. Dotted blue line: left stereo channel volume. Solid red line: right stereo channel volume.

The volume levels are readjusted every time the device detects a change in phone orientation or GPS location. The resulting effect is that the sound always appears to come from the direction of the target when it is toward one side of the user and at full volume from both speakers when the target is straight ahead. When the target is behind the user, the signal gets progressively weaker.

The sound played when the user sweeps over a POI can be described as a short beep. Different distance categories are represented with playing the beep sound with smaller pitch. Triggering the same target was limited to at most once per second in order to prevent fluctuating compass values triggering the same target rapidly.

Sound playback parameters used in the test were as follows. In both sweep and fetch features, targets in the "near" were considered to be the ones closer than 150 meters, "middle" targets closer than 1000 meters and "far" targets farther than that, up to 2000 meters. The width of the fetch cone was 30°.

The sweep feature imposes relatively strict requirements on the sound playback latency. In the ideal case, the user perceives the sound signal being played exactly when she/he his pointing the phone toward the POI in question. However, a number of factors induce latency in the sound playback: 1) the delay of magnetometer readings representing the changed environment upon turning the phone, 2) the delay of low-pass filtered accelerometer readings representing the actual phone orientation, 3) the time required for computing the appropriate sound pattern, and 4) the delay from the sound getting buffered for playback and it being actually played out from the speakers.

In Android API level 10, the sound playback feature was basically offered in the form of three different APIs, MediaPlayer, AudioTrack and SoundPool (which uses AudioTrack internally). None of these included proper support for low latency sound playback, which made factor 4 above the main problem in timing the sound for playback exactly when the user sweeps over a target. Initial tests were conducted to measure the playback latency. A sound effect was set for playback (using the MediaPlayer interface) upon release of a finger holding a button on the phone screen. The phone was dropped simultaneously 5mm on a desk, producing a soft thud. The timing difference between the thud and the sound effect was measured from a microphone recording. A small compensation for the falling time was added (calculated as 32ms for 5mm) to get the delay value. Repeating the measurement 10 times with Galaxy S and Nexus S (the models used in the field tests), the average total delay was 232ms.

Besides the problem related to accelerometer readings, the playback delay is another factor limiting the speed the phone can be swept when using the application. Experimental sweeping revealed a suitable maximum sweeping speed to be around 90° per second. This was measured by holding the phone in hand, rotating it at constant speed, and observing whether the compass tracked smoothly the correct compass direction.

4.6 Improvements

The tests described in this paper used the phone orientation method described in 3.2. However, gyroscopes have become more common after the tests were performed. The latest PING! version (implemented after the tests) uses Madgwick's orientation algorithm [4] that utilizes gyroscope, has low computational load and compensates for magnetic distortions. Although exact measurements have not yet been made, the first experiments with newer phone models equipped with gyroscope indicate that the new compass is more robust and allows for detecting the phone orientation with better precision when the phone is swung fast.

Moreover, newer phone models have shorter latency in sound playback. The playback latency measurements described above resulted for Samsung Galaxy Nexus the average total delay of 94ms and for Galaxy S3 161ms. Finally, the latest PING! version does not require a server, but can be used as a standalone application. POI sets can be disseminated as normal files.

5. EVALUATION

The application was tested in August 2011. The tests are discussed also in [3], with this paper providing additional details, in particular about the routes selected by the users and the difference between adult and teenager test users. The test group consisted of 24 people, 12 adults aged 21-50 (average 32) and 12 teenagers aged 12-15 (average 13,7). Of the adults, 3 were female and 9 male. All teenagers were female.

First, the users were introduced to the test and instructed in the use of the application for 30 minutes, including 10 minutes of actual hands-on instruction. After this, the users were sent to explore the city with the objective of locating three predetermined targets. Each group was given one mobile device connected to two stereo headphones. The users were given a generous limit of 90 minutes to complete the test. The three predetermined targets were a traffic sign in a particular corner, a hairdresser and the final meeting point. The locations were named in the application anonymously as “Target 1” etc. in order to prevent the users from utilizing their previous knowledge of the city. For checking whether the users had actually found the targets, the users were asked for a particular detail of the target (text written below the sign and the hairdresser’s name). After the test the users were interviewed and given a questionnaire to fill out.

Examples from the GPS track from four user test groups are shown in Figure 7. Most groups completed the test within 25-30 minutes, with some groups taking 50 minutes.

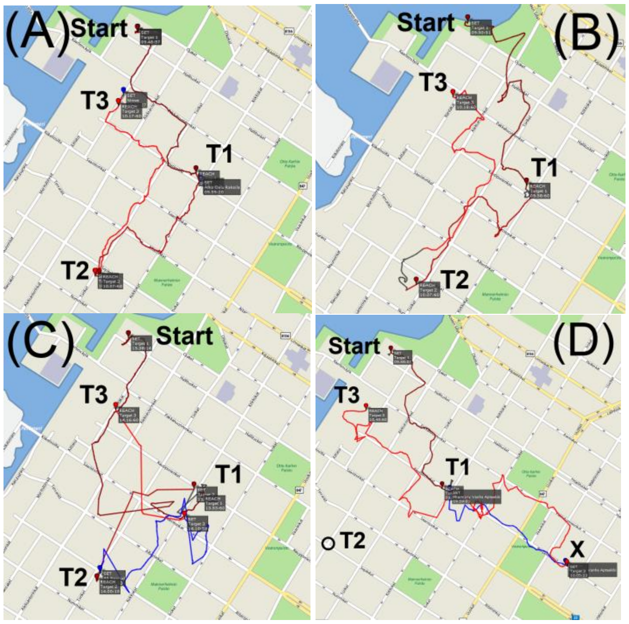

Figure 7. GPS tracks from four test user pairs. T1, T2 and T3 are target locations.

Of the 13 test groups, 10 succeeded in navigating the city and locating the targets without problems, for example as shown in A and B. Even though the GPS data in B is noticeably less accurate, both groups found the required targets. Two groups (one group shown in C) faced some problems with maintaining a GPS fix, while still managing to find the targets. One group lost the fix for several minutes due to taking shelter from rain in a coffee shop. One test group (shown in D) decided to search for an uninvolved location (X) instead of target T2.

In the interview, most users found the application sufficiently easy to use. The users appreciated the fact that the fuzzy routing lets them choose the final route themselves and clearly tells the correct direction without needing to actively orient to the surroundings or interpret a map. [3]. As a special mention, the users found the "target found" sound effect particularly satisfying, effectively conveying the idea of achievement.

The questionnaire was divided into two parts. The first part was based on the NASA Task Load Index (TLX) subjective workload assessment tool [9]. The original questionnaire was modified with an additional question “How much did the application hinder your awareness of the world around you”. The question about temporal demands was also removed because the test did not involve a tight time frame. Figure 8 shows the averaged results, with the answers ranging from “Very little” to “Very much” treated as numbers 1 (low) to 6 (high). Therefore the question about successfulness can be considered to have a positive result in the high end and the other questions in the low end. The results were split between the adults (left bar) and teenagers (right bar).

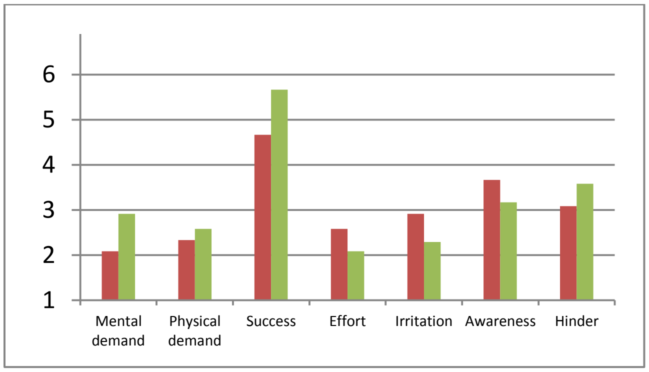

Figure 8. NASA TLX results.

The results indicate that the majority of the users felt success in navigating the city (5.2), and did not feel the application having significant mental (2.5) or physical demands (2.5) for successful use. The users did report the application causing to some degree reduced awareness of the surrounding world (3.3) [3]. However the teenagers reported being interested and positive about the application to noticeably better degree than the adults. This can be seen in Figure 8 with higher reported feeling of success (5.7), lower effort (2.1) and lower irritation (2.9), but also with lower awareness (3.2) and higher mental demand (2.9).

The second part of the questionnaire consisted of six statements regarding the navigation experience. The test users rated the statements using a six-level, forced choice Lickert scale, with the answers ranging from “Totally agree” (1) to “Don’t agree at all” (6). The statements are shown in Table 1, and the results in Figure 9. The number in each bar signifies the answer in question and the size the frequency. The statements for adults (A) and teenagers (T) are marked with a plus or minus sign based on whether agreement indicates a positive (+) or a negative result (-).

Table 1. Statements and their identifiers in figure 9.

ID	Statement
S1+	I felt secure and comfortable when using the PING application.
S2-	The PING application made me feel stressed while moving through the city.
S3+	While using PING I felt confident that the application would guide me correctly and help me find my way to the targets I selected.
S4-	When using PING, at several times, I felt lost and the application could not help me find what I was looking for.
S5+	I would like to explore other cities with the help of the PING application.
S6-	The sound of feedback from the PING application was frustrating and confusing.

Overall, when negative feedback was received, it was from users who had encountered a particular problem with the application, namely the loss of GPS fix or inaudible audio due to background noise. Feedback from other users was generally favorable. The statement results support the observation of the teenagers being considerably more positive about the application than the adults. This can be seen from the teenagers agreeing with statements S1, S3 and S5, and disagreeing with statements S2 and S4 considerably more than the adults.

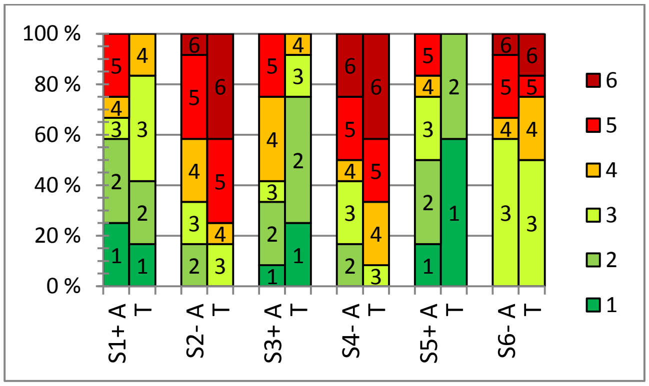

Figure 9. Questionnaire part 2 results.

The most striking difference can be seen with the teenagers' desire to explore other cities with the application, despite there being also users who reported not feeling confident about the application's ability to guide correctly. The results from statement S6 show over half of the test users reporting the sound feedback being frustrating and confusing to at least some degree. The primary complaints reported in the interview were that the sounds sounded too much like each other and the sound getting easily drowned by background noise [3]. Some users also reported difficulty with estimating the direction from the sound feedback. In these situations the users appreciated the option to fall back to following the graphical UI. On the other hand, the sound feedback was considered particularly useful when the phone display was difficult to see due to bright sunlight. This indicates that the sound feedback solution is still in need of further improvement, although it can be considered to have been of use for some users in the current form as well.

6. DISCUSSION

We presented a navigation application that lets users explore the POIs in their local environment by probing and guides users towards the selected POI with fuzzy routing: just providing the user with a direction and distance to the target, and delivered via minimal attention binaural audio and GUI. This lets the user decide the exact path to the destination her/himself while being kept oriented about the correct direction at all times, and without constantly studying the GUI. In contrast to traditional maps, there is no need to constantly relate the surroundings to the information on the map. The guidance is consistent regardless what route the user decides to take. This is in contrast to traditional navigation applications featuring turn-by-turn instructions where a previously given instruction, for example, to travel a certain distance may be contradicted by a new instruction in case the user misses a turn.

In the field test, the users succeeded in the given tasks and the application was commented to be easy to use. The tests, although focusing more on the guiding than the probing aspect, gave an initial verification to our hypotheses that PING! lets users focus on the environment. The test users valued the possibility to select the exact path themselves. As for the occasional problem situations, the tests indicated that especially with eyes-free UI, the status of the GPS fix should be conveyed more clearly to the user, for example with audio. The tests revealed a noticeable difference in attitude between adults and teenagers, but it is not yet known why exactly, since factors such as the fact that the teenager group consisted entirely of girls may have affected the result as well. Thus, more tests are needed, with varying test groups, in different environments and with more focus on the probing feature. The application also needs to be compared to alternative navigation applications.

The support for low latency audio in Android API level 10 and the used phone models was found inadequate. Considering a user swinging the phone at the estimated maximum of 90°/sec, disregarding other delays and taking only the delay between the audio command and the actual audio (estimated at 232ms) into account, the user has already rotated the phone about 20° by the time she/he hears the signal. In our experiments the actual delay was clearly noticeable. After the tests were conducted, the situation has in fact already improved somewhat with, recent more powerful phone models, gyroscope becoming more common, software libraries, and phone orientation algorithms [4], and can be expected to get better in the future. However, the problem with the sound playback latency is still a significant design issue when time-critical features such as the sweep action are concerned.

Besides focusing towards improving the hardware and algorithm performance, another approach to the latency problem described above could also be to measure the total delay, predict the motion of the user sweeping with the phone and time the sounds for playback in advance. Finally, the application could actively instruct the user to hold and swing the phone properly (i.e. slowly enough), which would help providing the user with accurate navigational instructions even when modern orientation sensors are unavailable. Considering the recent developments and the fact that the user tests described in this paper focused more on the guiding aspect than the sweep feature, further tests are needed in order to understand the potential of the sweep feature.

As the application requires listening to audio from stereo headphones, two challenges need to be considered: the application blocking the ambient soundscape and the ambient soundscape blocking the application sounds. The first problem is partially mitigated by the fact that PING! is designed to work from background – the application stays mostly silent except when it provides the occasional guiding audio signals. However, in a noisy environment the user might not hear the audio signals well enough to successfully estimate the direction to the target. This might encourage him/her to use audio blocking headphones that end up blocking the ambient soundscape even when the application is silent. This might lead to reduced awareness of the surroundings and danger – especially with traffic.

One solution could be to replace the binaural audio signals and stereo headphones with mono audio played from the phone’s speakers, which simply tells the user if the navigation target is inside a pre-determined cone reaching out from the top of the phone. The work by Magnusson et al. [14] investigates various widths for such a cone. Ambient soundscape could also be integrated in the application and played through the earphones – either just attenuating the volume, or analyzing the soundscape and extracting only the sounds recognized as important. One more option would be to use bone conduction headphones (see, e.g. www.chilli-tech.com) that would potentially let a user be aware of the environment, while also able to hear the guiding sounds.

The fuzzy routing method described in this paper needs to be improved to handle better situations when the traversable path differs considerably from the direct path to the target, for example, due to a large obstacle. This may happen especially at longer distances, in city environments and close to natural obstacles. An example with a river blocking the shortest path is illustrated in Figure 10 (left). In case like this the way across the obstacle may not be immediately obvious, so the application should detect the situation and direct the user towards a better route (here: via the bridge). One solution for the problem would be to introduce a system for routing the user around such obstacles using waypoints. In the spirit of fuzzy routing, the number of waypoints should be kept as low as possible. This could be achieved by first calculating a turn-by-turn route using any route service providing point-by-point routes in an open format and then filtering the route so that a minimal number of waypoints are left. An example of filtered route with a minimal number of waypoints for going around an obstacle is shown in Figure 10 (right).

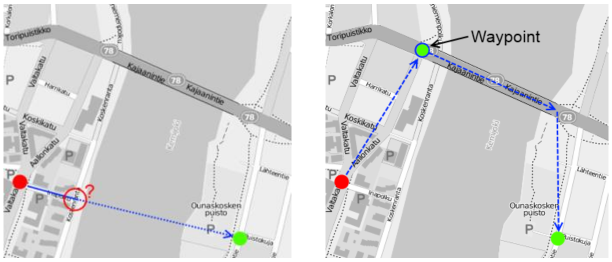

Figure 10. Left: Bad routing through a natural obstacle. Right: Routing around a natural obstacle with algorithm-placed waypoint at the beginning of the bridge.

Finally, although the parameter values determining sweep and fetch behavior (30° cone) worked relatively well during the tests, they are not yet optimized. The fixed values may also result in a situation where the scan results in too many results in high POI density areas, or too few results in low density areas. Here, the application would benefit from adapting the parameter values to the environment instead of fixed values.

All in all, although there is still room for improvement, the field tests indicate that we are on the right track. The testers clearly expressed the desire to explore unknown cities with this type of application. When the amount of interaction and focus of attention is considered, the requirements are small. With the current interface, from starting screen the user needs four button clicks in order to start navigation. In particular, the user does not need to study maps or write text in order to search for targets in the vicinity, as searching is done by sweeping the phone. This can be seen of particular importance when the user does not have much previous knowledge about the surroundings. In this case the user can just point the phone to a direction for immediate feedback about the benefits of proceeding to that direction and enjoy the surroundings while navigating. This also matches well with the concept of calm computing [15]: computers vanish in the background and support us in our everyday activities without demanding too much focus or disrupting our activities.

When comparing the data required for navigating using PING! to contemporary navigation applications, the most notable difference is that PING! requires no data regarding roads, speed limits, one-way streets, etc. This means that the database is relatively small and could easily be downloaded on-demand. Also, contemporary navigation applications most often use proprietary navigation databases. Because PING! only needs information about possible navigation targets, it could use a community driven database which could also be free, or relatively cheap, to use. In fact, the latest version completed after the field tests supports exporting and disseminating POI sets as normal files. Moreover, the lack of continuous connection to a database on a server decreases the amount of communication and the latencies – and makes the software simpler and more robust.

Generally, this kind of navigation would be a good addition to a future mobile device emphasizing user interfaces that minimize distractions to everyday activities. Such a device could have binaural audio feedback, speech recognition, and support for tangible user interfaces. Apple's Siri illustrates how audio-based user interfaces could be used much more, and we have in our earlier work demonstrated the potential of tangible user interfaces, specifically NFC-based ones [16]. In fact, users might carry one device that is comfortable in hand, presents crucial information (like the arrow in PING!) and can be used to point, touch, and to perform other physical actions to interact with the applications. Then, another pad-like device could offer large and high-resolution display for graphical and textual content.

7. CONCLUSIONS

The PING! navigation application allows users to probe points of interest by sweeping the phone and listening to audio feedback. The application does not present maps or exact turn-by-turn directions but utilizes fuzzy routing and indicates the direction to the target with a simple arrow and binaural audio. The conducted field tests supported the feasibility of this kind of application concept, with the test users reporting interest towards using the application for exploring new environments. The tests also revealed ways to improve the application further, especially by helping the user cope with problem situations, such as a lost GPS fix. All in all, we are confident that this kind of applications with modest GUI requirements possess a significant potential and can play even a major role in the everyday life of future mobile users.

8. ACKNOWLEDGMENTS

The II City project was funded by EU Interreg IV A North, the County Administrative Board of Norrbotten (Sweden), the Regional Council of Lapland (Finland), and The Interactive Institute (Sweden). Mikko Pyykkönen and Jari Koliseva from University of Oulu, Mats Liljedahl, Katarina Delsing, Nigel Papworth and Stefan Lindberg from the Interactive Institute, and Ismo Alakärppä and Elisa Jaakkola from University of Lapland are acknowledged for their participation to the work reported in this paper.

9. REFERENCES

[1] Magnusson, C., Breidegard, B. and Rassmus-Gröhn K. 2009 Soundcrumbs – Hansel and Gretel in the 21st century. *HAID 09 (4th International Workshop, Haptic and Audio Interaction Design)* 10.-11. Sept. 2009 Poster.

[2] Holland, S., Morse, D. R. and Gedenryd, H. 2002. AudioGPS: Spatial Audio Navigation with a Minimal Attention Interface *Personal and Ubiquitous Computing*. 6,4 (2002), 253-259

[3] Liljedahl, M., Lindberg, S., Delsing, K., Polojärvi, M., Saloranta, T. and Alakärppä, I. 2012 "Testing Two Tools for Multimodal Navigation," Advances in Human-Computer Interaction, vol. 2012, Article ID 251384, 10 pages, 2012. doi:10.1155/2012/251384.

[4] Madgwick, S. O. H., Harrison, A. J. L. and Vaidyanathan, R. 2011 Estimation of IMU and MARG orientation using a gradient descent algorithm *2011 IEEE International Conference on Rehabilitation Robotics*, June 29th – July 1st 2011, pp. 1-7. doi: 10.1109/ICORR.2011.5975346

[5] Motion Sensors (Visited Sept 2nd 2012) http://developer.android.com/guide/topics/sensors/sensors_motion.html

[6] Ochieng, Washington Y., M. Quddus, and Robert B. Noland 2003. "Map-matching in complex urban road networks." *Brazilian Journal of Cartography no.* 2, vol. 55, pp 1-14

[7] Mazhelis, O. 2012 Using recursive Bayesian estimation for matching GPS measurements to imperfect road network data. *13th Intl IEEE Conference on Intelligent Transportation Systems* (ITSC), 19.-22. Sept. 2010. pp.1492-1497, doi: 10.1109/ITSC.2010.5625138

[8] Liljedahl, M. and Lindberg, S. 2011 Sound parameters for expressing geographic distance in a mobile navigation application. *Proceedings of the 6th Audio Mostly Conference: A Conference on Interaction with Sound.* pp. 1-7, http://dx.doi.org/10.1145/2095667.2095668

[9] NASA TLX: Task Load Index (Visited Feb 20th 2013) http://humansystems.arc.nasa.gov/groups/TLX/index.html

[10] Pascoe, J., Ryan, N. and Morse, D. 1999 Issues in developing context-aware computing. In: Gellersen H-W (ed). *First International Symposium on Handheld and Ubiquitous Computing* (HUC 99): Lecture Notes in Computer Science 1707. Springer, Berlin Heidelberg New York, 208-221.

[11] Strachan, S., Eslambolchilar, P., Murray-Smith, R., Hughes, S. and O'Modhrain, S. 2005 GpsTunes: controlling navigation via audio feedback. *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services* (MobileHCI '05) pp. 275-278, doi: 10.1145/1085777.1085831

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

[13] Garmin Forerunner 410 sports watch. (Visited Sept 3rd 2012) https://buy.garmin.com/shop/shop.do?cID=142&pID=83274

[14] Magnusson, C., Rassmus-Gröhn, K. and Szymczak, D. 2010 Show me the direction – how accurate does it have to be? *Using Audio and Haptics for Delivering Spatial Information via Mobile Devices, Workshop at MobileHCI 2010*, 7. Sept 2010. pp. 22-25

[15] Weiser, M., Gold, R., and Brown, J.S. The origins of ubiquitous computing research at PARC in the late 1980s. In: IBM Systems Journal, vol. 38, no. 4, pp. 693 - 696, (1999) ISSN: 0018-8670, doi: 10.1147/sj.384.0693

[16] Riekki, J., Sánchez, I. and Pyykkönen, M. 2012 NFC-Based User Interfaces. *4th International Workshop on Near Field Communication*, March 13th 2012, pp. 3-9, doi: 10.1109/NFC.2012.19