Timbremap: Enabling the Visually-Impaired to Use Maps on Touch-Enabled Devices

Jing Su
Dept of Computer Science
University of Toronto

Alyssa Rosenzweig
Dept of Computer Science
University of Toronto

Ashvin Goel
Dept of Electrical and
Computer Engineering
University of Toronto

Eyal de Lara
Dept of Computer Science
University of Toronto

Khai N. Truong
Dept of Computer Science
University of Toronto

ABSTRACT

Mapping applications on mobile devices have gained widespread popularity as a means for enhancing user mobility and ability to explore new locations and venues. Visually impaired users currently rely on computer text-to-speech or human-spoken descriptions of maps and indoor spaces. Unfortunately, speech-based descriptions are limited in their ability to succinctly convey complex layouts or spacial positioning.

This paper presents Timbremap, a sonification interface enabling visually impaired users to explore complex indoor layouts using off-the-shelf touch-screen mobile devices. This is achieved using audio feedback to guide the user's finger on the device's touch interface to convey geometry. Our user-study evaluation shows Timbremap is effective in conveying non-trivial geometry and enabling visually impaired users to explore indoor layouts.

1. INTRODUCTION

Mapping applications on mobile devices have gained widespread popularity as a means for enhancing user mobility and ability to explore new locations and venues. These applications can provide two features, *navigation* to a specific location and *exploration* of an area to obtain a sense of the general layout of streets and locations. For visually impaired users, current mapping applications are mostly text-to-speech (TTS) driven, whereby lists of directions or items are enumerated to the user. These applications are primarily for real-time navigation, providing turn-by-turn walking directions [9, 16].

Unfortunately, TTS is poorly matched for map and geometry exploration due to the difficulty in producing clear and succinct descriptions of complex shapes and positions of features. This shortcoming is apparent even in human conversation when attempting to describe a complex layout. In formative interviews that we conducted with visually impaired participants, several have indicated that a common technique for addressing this limitation is to "draw" with their finger on the other person's hand, illustrating a shape or position through tactile interaction and feedback.

A second significant limitation of existing mapping applications is that they mainly support public outdoor environments. There is no pervasively available localization system for indoor environments, such as office spaces and large malls, to aid in navigation or feature searching. Indoor environments are particularly difficult to explore or navigate for the visually impaired. While some buildings have maps and diagrams to show the layout of the indoor space, these diagrams are rarely accessible to the visually impaired. Furthermore, even at locations with braille or tactile maps, it can be difficult for individuals to determine the correct orientation and direction of travel [2]. More commonly, visually impaired users will call accessibility hotlines or ask for help from strangers.

This paper introduces Timbremap, a novel interface for map and floor-plan exploration. Timbremap utilizes a touch surface to convey complex geometrical information. The interface provides output feedback in the form of sonification (non-speech audio to convey or perceptualize data) to guide the user's finger along shapes, enabling the user to develop a cognitive understanding of the paths and features within a layout.

We implemented and evaluated Timbremap on a touch-based off-the-shelf mobile phone, enabling users to explore maps and floor-plans both at home and away. Timbremap enables users to contextualize the world around them, such as when entering an unfamiliar hotel, office space, or shopping mall. Furthermore, Timbremap on a mobile platform enables users to deal with unexpected situations and plan changes, and reduces the cognitive load associated with memorizing floor-plans.

The design of Timbremap presents two challenges. First, the interface must be effective in guiding the user's finger on a smooth touch-screen surface. Second, the interface must function within the confined surface area of a mobile phone.

This paper presents two major contributions. First, we present an implementation of the Timbremap interface on an off-the-shelf mobile device. We present two sonification modes – one actively guides the user's touch, based on the underlying layout information, while the other uses a graph-coloring technique to passively convey information at the touch point. Our user-study evaluation shows Timbremap is effective in conveying complex geometry information to visually impaired users. Our participants achieve an average accuracy of 81% in shape identification, spending just over 41 seconds per shape. Second, we show that the Timbremap

application is effective in conveying complex indoor floor-plans. A user is able to build a mental map of a complex indoor floor-plan, describe the points of interest, and provide directions for navigating around the indoor space.

These results are highly encouraging, thereby motivating future development of mobile and ubiquitous platforms for providing accessible exploration of indoor and outdoor map and location information. Furthermore, there is significant potential in expanding Timbremap's interface to medium and large format touch devices such as tablets and tabletops.

2. BACKGROUND & RELATED WORK

This section covers three major areas of work. First this section will describe works exploring techniques for contextual cue and notification systems. Next, we cover works exploring methods for conveying geometry. Finally, while Timbremap focuses primarily on exploration rather than navigation, we briefly address several works which investigate techniques for providing visually impaired users with spacial context and navigational aid.

2.1 Non-visual notification

Many research efforts have examined audio and haptic techniques for conveying eyes-free notification and context information. Automated aids for user notification are an important line of investigation, even if tactile notices, such as braille tags, are available. A difficulty with physically placed tags is that the user often cannot know a priori where the tags will be or how they are oriented [13], especially in unfamiliar environments.

Dingler et al [4] examined various sonification techniques for compressed audio notification. The motivation is to create the audio equivalents of signs and icons. Audio icons, or earcons, were evaluated to explore their effectiveness in conveying metaphoric meanings, for example an ascending tri-tone conveys "up". Dingler et al also explored the use of spearcons, which are highly compressed, short sequences of speech. They found that spearcons were highly effective in conveying the spoken meaning to the user while not imposing the cognitive load that standard speech incurs on the human listener.

The AUDIOGRAPH system [1] explored enhancements to the earcon concept, whereby musical sequences or relationships between musical sequences convey semantic information. Their results show that sophisticated audio sequences can be used to convey complex data, but users need to have understanding of musical concepts and be trained in the interface and musical concepts.

SemFeel [17] uses high-precision actuators to provide complex tactile feedback sensations. While this technique is effective at providing general notification feedback, it is currently not sufficient for conveying high-precision geometry. Furthermore, SemFeel requires complex add-on hardware that is not readily available as an off-the-shelf consumer product.

Finally, several efforts have investigated techniques for eyes-free menu and item selection on mobile devices [18, 8]. In particular, Timbremap utilized some techniques presented in Slide rule [7] for item selection on a multi-touch mobile interface.

2.2 Geometry and spacial exploration

Jacobson [6] utilized a touch-pad to convey relative positioning of points of interest on a map using sonification. Timbremap extends this effort by showing effectiveness on a constrained mobile device and providing sufficient precision to determine complex geometric shapes.

Other projects have utilized tactile methods for exploring and conveying geometry. The BATS [12] project used force-feedback joysticks coupled to a pointer for providing tactile bumps and feedback over an interface as the cursor crossed boundaries and feature changes. Crossan et al [3] used a force-feedback 3-dimensional pen to guide the user's hand in a trajectory, outlining a geometry. These methods suffer from requiring expensive or non-portable add-on hardware. In contrast, Timbremap is designed to work with off-the-shelf mobile devices without need for complex add-ons. Furthermore, Crossan et al's results [3] show that even with a highly precise tactile or force-feedback system, audio cues were still necessary. This suggests that audio systems are likely to be needed to complement interfaces for the visually impaired, even if highly precise deformable or tactile surfaces are developed.

2.3 Navigation

While the current focus of Timbremap is not as a navigational assist, these related works are worth discussion as avenues for expansion and integration.

The SWAN [16] project explored the use of non-verbal auditory cues (*i.e.*, sonification) to aid visually impaired users in way-finding. Unlike current GPS road-maps which typically guide users based on streets, the SWAN system can potentially guide users along way-points across non-road areas such as parks or large campuses. The approach used in the SWAN project is to use sonar-like beacons in a spacial audio system to guide users from way-point to way-point. Other efforts [10, 5] examined the use of a remote which can be used to point at and query for environmental objects and information – effectively providing a "smart" enhanced cane. Marston et al [11] compared a haptic and audio based navigation systems, using in situ navigation in both a small city block as well as a park.

Subsequent follow-up work [14] examined the impact of users navigating via audio, while at the same time discriminating a secondary audio announced task (such as carrying a conversation). Their results show that contrary to initial assumptions, navigation performance degraded when presented with multiple audio demands, instead of the secondary audio task. This result is a more minor concern for the Timbremap interface, since we expect users to stand still and utilize it to explore or reference their location, rather than use it as a navigational aid.

3. TIMBREMAP INTERFACE

The primary objective of the Timbremap interface is to provide a clear method of conveying non-trivial geometry, spacial positioning, and feature information by guiding the user's touch exploration of the device's touch surface. To this end, we have developed an interface which can handle a wide range of complex geometries found in most roads and indoor spaces. Because the Timbremap interface focuses on the use of audio feedback, we decided to constrain the interface design to exploration so as to not potentially overload or confuse an actively navigating user.

Through an iterative design process we present and evaluate two sonification modes for the Timbremap

interface. First we present a pro-active sonification mode which attempts to guide the user's touch. Then we present a passive sonification mode which conveys spatially connected features. Finally we present an indoor floor-plan exploration application using the Timbremap interface, and its implementation on an iPhone device.

3.1 Line hinting mode

The Timbremap *line hinting* sonification mode provides pro-active feedback to help guide the user's touch within the path segments of a layout. Should the user's touch point wander away from the path segment, hinting sounds are played to guide the user toward the nearest path segment, relative to the touch point.

The design process started with an initial interface which provided audio feedback when the user's finger was over a path portion of the shape. The audio feedback pans and fades as the user's finger wanders away from the path. Early user feedback on this interface suggested that the audio feedback did not provide sufficient information to what corrective action the user should have taken. Stereo positioning of the path sound, attenuated to the relative deviation of the user's finger to the path, was found to be too subtle.

To address this issue, we implemented an audio hinting system which indicated to the user what corrective action to take in order to return to the nearest path point. The first iteration of the hinting system used mid-tone beeps on left/right stereo channels, and up/down were indicated using high and low tones, respectively. User feedback indicated that the high and low pitched tones for up/down were too abstract. To address this, we replaced the high/low tones with high and low-pitched spearcons [4, 15] (compressed speech audio icons). Furthermore, the initial iteration supported diagonal hints – for example playing a high-pitched tone in the left ear to indicate upper left. Many users found the diagonal hints to be too complex. We subsequently restricted this hinting system to the four compass directions during the experiment.

In designing the sonification for the shapes and paths, we found a trade-off had to be made between the width of the path segments and the sensitivity of the interface. Making the path segments very thin enabled users to detect subtle changes such as minor curves. However, the thin segments were also much less forgiving, resulting in "noisy" feedback, constantly hinting the user towards the thin line. Conversely, a thicker line was found to be much more relaxing and forgiving to use, but made subtle curves difficult to detect. To address this, we use a slightly modified sonification for curved segments. The curve sound does not provide information as to the direction or degree of the curve. This limitation on curve sonification was designed for two reasons. First, it reduces the amount and complexity of audio feedback the user must process. Second, interviews with participants revealed that how a curve should be sonified depends on the user's perception. For example, a curve perceived as turning to the right is perceived as turning to the left if travelling back in the reverse direction.

Finally, we added sonification markers for intersections. We defined an intersection as any point where three or more segments connect. Thus, a right-angle turn is not considered an intersection, but a T-junction is. The intersection sound not only helped notify the user of an intersection, but also
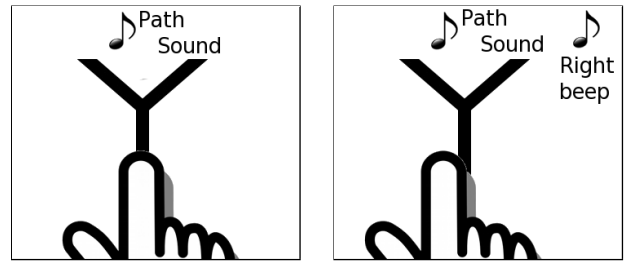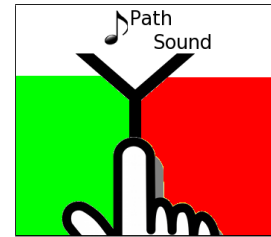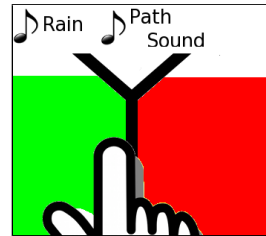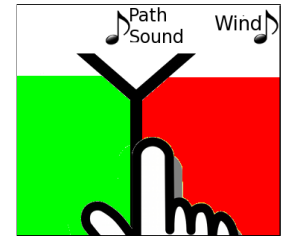

Figure 1: Line hinting example. As finger shifts to the left, sonification beeps on the right to indicate path position

(a) finger on path; path sound is played

(b) finger shifted to left; sound for left side coloring and faded path sound is played

(c) finger shifted to right; sound for right side coloring and faded path sound is played

Figure 2: Area mode

serves as a reference point.

Figure 1 illustrates an example. First, the user hears a repeating audio pattern representing the path segment. As the user's touch shifts to the left, a repeating beep in the right ear is played to indicate that the nearest segment is to the right of the finger. This interface uses mid-tone beeps on stereo channels to indicate whether path is to the left or right of the touch point. For up and down adjustments, this uses high-pitched "up" and low-pitched "down" spearcons, respectively. With practice, users should be able to respond very quickly to the audio hints and follow path segments accurately.

3.2 Area hinting

Observation of user strategies in the iterative design process of the line hinting mode led us to develop an alternative sonification, which we called the *area hinting* sonification mode. We found that users often swept their finger across the screen surface to pick up global characteristics such as number of segments around edges, gaps between segments, and existence of any intersections. The area hinting mode had three design goals: (1)

remove explicit directional sonification while providing left/right/up/down information; (2) enable users to deduce contiguous regions of empty space; and (3) improve clarity of feature detection when using a swiping strategy.

As with line hinting, the area hinting mode uses the same sonification strategy for denoting straight and curved paths and intersections. The blank spaces surrounding the paths are flood-filled with sound markers, assigned using a 4-coloring heuristic. These four markers are mapped to a set of four ambient sounds: light raindrops, wind, chirping crickets, and gentle wind chimes. This sonification system enables users to detect whether they are straying from a path segment, and in what direction, by listening for and comparing changes in sound. For any given segment, the ambient sound for either side of the path will be different. An example of the audio feedback is illustrated in Figure 2. On the path, the user hears a repeating audio pattern. If the user's finger shifts to the left, the path sound fades out and the sound of rain fades in from the left. If the user's finger shifts to the right, the sound of wind fades in from the right. By swiping over the touch surface, users can assess the number of segments as well as the connectivity of empty spaces.

3.3 Indoor exploration application

Using the sonification described above, we implemented a basic indoor floor-plan exploration application. In addition to providing sonification for paths, spaces, and intersections, this application added a sonification for marking points of interest (POI). Identifying POI markers and map panning is supported via intuitive multi-touch interactions [7]. Specifically, to identify a POI marker, the user holds one finger on the POI marker, and double-taps anywhere on the screen with a secondary finger. This commands the system to read the marker using a TTS voice.

To pan the map, the user first positions their primary finger on any spot on the map. With a secondary finger, the user holds any of the four corners of the screen. By dragging the primary finger, the map pans by moving with the primary finger. Figure 3 illustrates an example map panning operation, to pan the "stairs down" point of the map. With a finger over the "stairs down" marker, the user places a second finger on one of the corners, in this case the top left corner. Then as the user pans the primary finger, the map pans with the finger. By holding and releasing the corner touch, the user is able to quickly "walk" and pan across multiple screens without losing their place.

3.4 Device

We chose to use the iPhone as the mobile device platform for two reasons. First, at the time, the iPhone was the only available mobile smartphone which featured a capacitive touch-screen with support for multiple simultaneous touches on the screen. We chose a capacitive screen instead of a resistive screen because we felt it would be a more ergonomic touch mode when primarily using fingers. The actual device used was an iPhone 3G, running the 3.0 firmware. The application is implemented as a native iPhone application, written in Objective-C. Storage for reading shapes and logging data are done using the SQLite3 library provided as part of the iPhone application API.

For aesthetics, the iPhone is designed with many smooth surfaces. While visually appealing, this results in a device

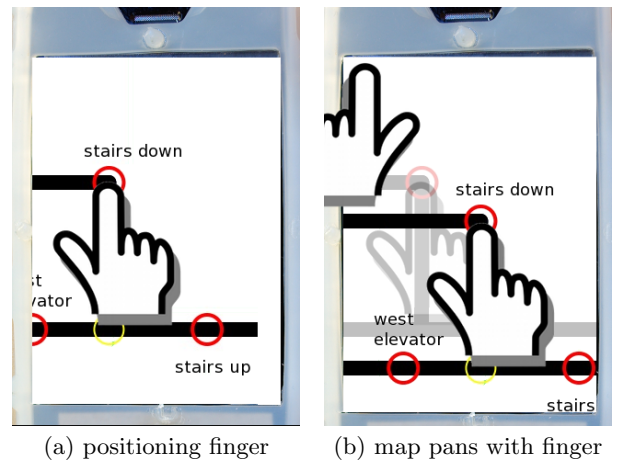(a) positioning finger   (b) map pans with finger

Figure 3: Panning operation. By holding any of the corners, the user can drag the map.

which can be slippery to hold and lacking in tactile feedback for certain physical features. To mitigate this, we wrapped the iPhone in an after-market silicon "glove". The silicon border enabled users to feel the boundaries of the touch screen area. This was important since the top and bottom edges of the touch screen area are smooth with respect to the earpiece and home button areas. The non-slip silicon glove also enables users to use both hands to interact with the iPhone on a table-top without the device sliding around. Figure 4 shows the iPhone device in its silicon glove.

In this prototype implementation, we assumed users will use stereo headphones in order to take advantage of the stereo sonification techniques Timbremap employs. Closely positioned or monophonic on-device speakers may not provide sufficient left and right channel distinction. Since Timbremap is focused on exploration rather than navigation, we do not expect the use of headphones to negatively impact visually impaired users.

4. EXPERIMENTAL SETUP

The objective of the experiment and user study is to first determine whether the use of sound hints combined with touch interaction in the Timbremap interface is effective in conveying geometry information; second, determine whether there is a significant difference in the effectiveness of the *line* and *area* hinting modes; and third, whether these basic skills translate to an ability to explore an indoor floor-plan.

4.1 Shape identification and discrimination

This first user study is designed to quantitatively determine if users are able to identify and distinguish between different shapes using the two sonification modes described in the previous section. We begin with a description of the experimental procedure in this user study, followed by a description of the shapes and selection tiles in the following subsections. Participants will use the prototype interface implemented on the iPhone device, along with a pair of stereo headphones for the sonification.

This study is divided into three parts: a tutorial, practice, and test. In the tutorial segment, the participant is given time to be acquainted with the interface and sonification
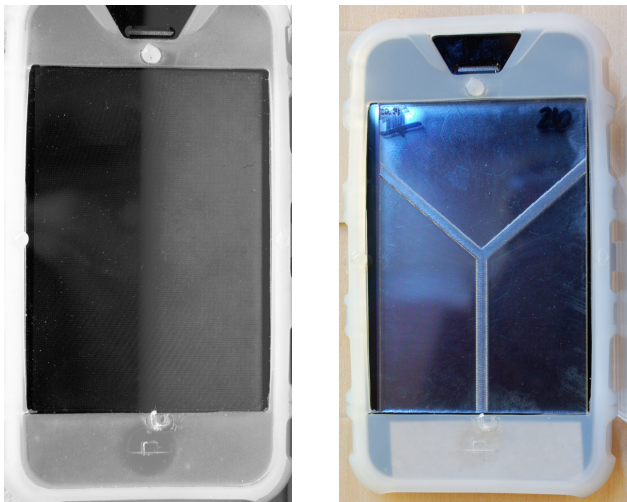
(a) iPhone in silicon glove

(b) tile dimensions match device's screen dimensions

Figure 4: iPhone and sample tile.

Figure 5: Participant using the Timbremap device during a study. The shape is shown on the device screen. This particular participant is completely blind and could not see the shape. A Y-splitter was used in the audio jack to enable the experimenters to also hear the sonification.

using sample shapes which do not appear in the test portion. In the practice segment, participants were given eight practice shapes. In the first four practice shapes, the participant is first presented with a physical etched tile of the shape (described later in Section 4.1.1), and then asked to look for the shape on the device's interface. In the later four practice shapes, the participant is presented with the device's interface first, and then asked to find the shape amongst a set of possible tile choices. This practice portion is not timed, and participants are given immediate feedback on the correctness of their selection. At the end of the practice session, participants are given the option to retry any of the practice shapes until they are comfortable with the workings of the system.

The test portion of each interface is divided into two blocks. Each block presents 24 shapes, including repeated presentations. These 24 presentations are chosen from coreshapes and distractors, described in Section 4.1.1. We present 18 coreshapes (three repeats of six coreshapes), and 6 distractors. The goal of this arrangement is to determine if participants are able to detect general properties of the shapes, as well as differences between similar shapes. By using two test blocks, we are able to test if there is a noticeable learning effect.

For each shape, the participant is first presented with the device interface, and given at most 60 seconds to interact with the interface, as illustrated in Figure 5. Participants are encouraged to finish before the expiration of the timer if they are confident they know what shape is being presented. The participant is then given 30 seconds to select from amongst three engraved tiles. These three choices are placed in a wooden holder, as shown in Figure 7b. One tile choice is an exact match to the displayed shape, with the two others being similar in some features. We use a multiple choice answer selection format to remove ambiguity that might otherwise result if participants were asked to describe or draw the displayed shape. After each block
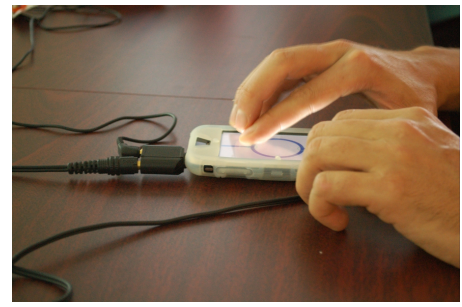
of 24 presentations, the participants responded to a brief questionnaire and given a short break before repeating the three-part process again with the other hinting sonification mode.

4.1.1 Shapes

One goal of the shape design process was to select a variety of shapes which conveyed different road or indoor path features, in incremental degrees of complexity. Thus, we were not interested in overly arbitrary or abstract geometries, but rather to determine participants' ability to distinguish between realistic shapes. Figure 6 illustrates the final set of shapes used in the test portion of the experiment. These shapes are divided into two main categories: *coreshapes* and *distractors*. The coreshapes, as seen in Figure 6a are similar in concept, each designed with a subtle variation and increasing in complexity. The distractor shapes (Figure 6b) are designed to be a departure from the coreshape variations and prevent participants from making a choice based on simple characteristics such as intersection count or number of bordering segments.

4.1.2 Tiles

To provide a real-life tactile representation of these shapes for the visually impaired participants, we chose to engrave them onto acrylic tiles. These tiles are cut to the same dimensions as the iPhone device used as the mobile platform. The widths of the engravings match the widths as presented on the device's interface. In designing the engraved tiles, we decided to engrave the path and leave the "blank" areas smooth, as opposed to raising the path above the surface. This was decided for two reasons. First, the grooved etch along the path provides a rougher, stronger tactile sensation. This positive feedback reinforced the presence of the shape, while the smoothness of the blank areas conveyed a sense of "nothing". Furthermore, the etching produced a groove which helped guide the participants' finger along the path. A close-up picture of an index finger along a tile's groove is illustrated in Figure 7a.

For participants to feel and compare multiple tiles, we used a wooden frame which can hold three interchangeable tiles. This frame, illustrated in Figure 7b with three sample tiles, provides a convenient method for participants to find and compare three given choices.
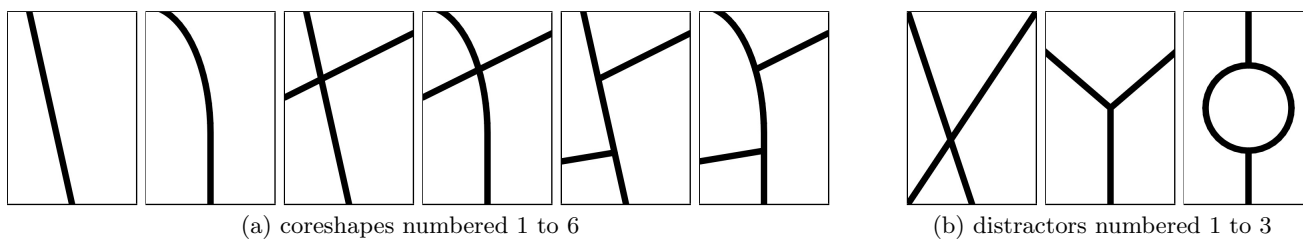
(a) coreshapes numbered 1 to 6

(b) distractors numbered 1 to 3

Figure 6: Shapes used in experiment. Coreshapes are incremental in complexity. Distractors used to prevent educated guessing

(a) groove dimensions relative to index finger
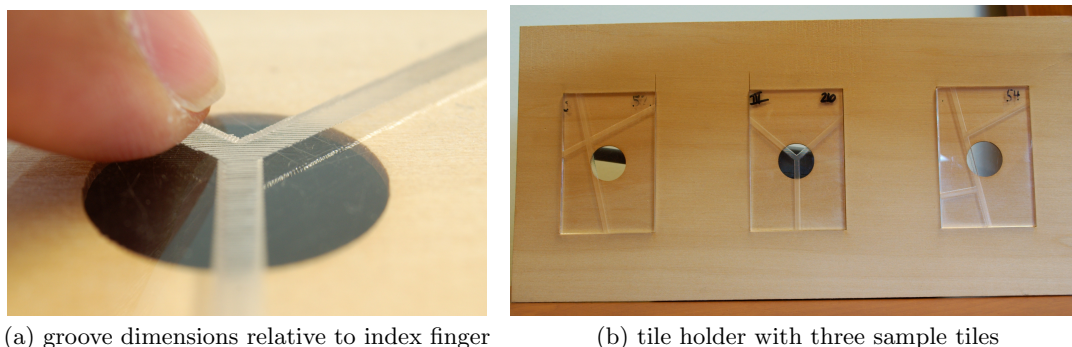

(b) tile holder with three sample tiles

Figure 7: Close-up of tile groove and multiple-choice wooden tile holder.

4.2 Indoor exploration application

In a second user study, we qualitatively examine whether the shape identification and discrimination skills utilized in the first user study translates to an ability to pan a large indoor floor-plan and build a cognitive map of the space and its features.

Two maps were supplied, one more complex than the other. Each map is larger than can be displayed in one screenful, and contains several POI markers. An example indoor floor-plan is shown in Figure 8. This study imposed no time limits. The participant was given as much time as he liked to explore the map until he was comfortable and confident of the layout. The participant was then asked several point of interest and walking direction questions.

4.3 Participant pool

Our first user study consisted of 6 individual participants; 2 female and 4 male. The participants for the user study were separate from the participants in the iterative design process. Three participants reported their age in the 30's; one participant in their 40's; one participant in the 50's; and one participant in their 60's. We had one congenitally blind participant, and all others completely or nearly blind for at least 10 years, save for basic light sensitivity. Four participants relied primarily on a guide-dog for navigation, and two relied primarily on a cane. The second user study used one returning participant from the first study.

All of the participants had at least some braille reading ability. On a five-point scale (with five being excellent), two participants scored themselves as 4 and 5, with the remainder scoring themselves as either 1 or 2. All participants reported having extensive experience or familiarity with mobile phones and computers; only two participants have technical occupations relating to

computers and software. None of the participants were affiliated with the university or research lab.

Participants were recruited via a mailing list of local and diverse visually impaired users interested in research participation. Participants were scheduled for a 3-hour block of time, and granted a $60 honorarium for their time.
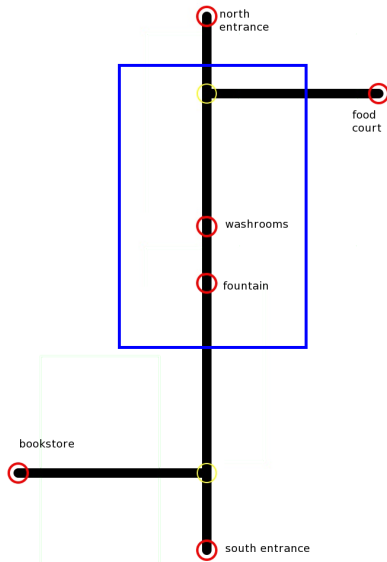
5. RESULTS

The following two sections describe the results of our two user studies.

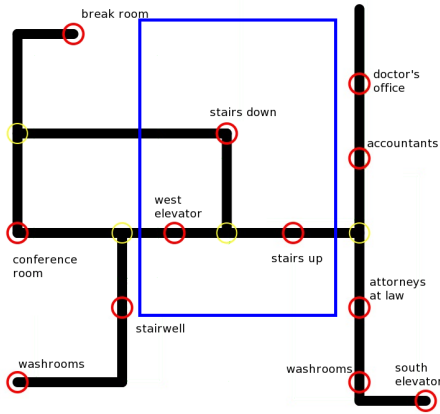## 5.1 Shape identification and discrimination

We carried out the experiment using file trials of the line interface and five trials of the area interface. Four out of the six participants were able to try both the line and area hinting sonification modes. We counter-balanced the presentation of the two sonification modes to eliminate learning bias.

We first examine the correctness results of participant answers, detailed in Table 1. A summary of the mean accuracy with standard deviations between the distractors, core shapes, and all shapes combined, is shown in Figure 9. Across both modes, participants were very successful at determining the shape within the time constraint, averaging over 80% accuracy.

We examined whether there is an accuracy difference between the two modes, and whether there is an improvement due to learning from block 1 to block 2. We found no statistically significant difference in the means and deviations between the line and area hinting sonification modes ($T = 0.097 < T_{crit}^{\alpha=0.05} = 2.440$). This suggests that both modes are effective. Based on participant feedback, the choice of which to use should be left to user preference. Future work may examine whether users have

(a) map 1

(b) map 2

Figure 8: Indoor floor-plans. POI and intersection markers indicated by labelled and unlabelled circles, respectively. Size of device screen coverage shown by blue box.

	block1	block2	combined
distractor 1	90.0	83.3	85.0
distractor 2	90.0	66.7	85.0
distractor 3	70.0	66.7	65.0
coreshape 1	100.0	100.0	100.0
coreshape 2	93.3	100.0	96.7
coreshape 3	86.7	88.9	86.7
coreshape 4	70.0	77.8	70.0
coreshape 5	66.7	77.8	63.3
coreshape 6	70.0	88.9	70.0

(a) line hinting sonification

	block1	block2	combined
distractor 1	83.3	66.7	75.0
distractor 2	83.3	83.3	83.3
distractor 3	50.0	83.3	66.7
coreshape 1	88.9	100.0	94.4
coreshape 2	100.0	88.9	94.4
coreshape 3	100.0	77.8	88.9
coreshape 4	66.7	100.0	83.3
coreshape 5	88.9	55.6	72.2
coreshape 6	77.8	55.6	65.6

(b) area hinting sonification

Table 1: Aggregate percentage answered correctly

of fatigue and test for possible improvements due to learning.

Next, we examined the amount of time that participants needed to determine a correct answer. On average, participants were able to determine the correct shape after approximately 41 seconds. The full findings are detailed in Table 2. A summary of the distribution of mean completion times is summarized in Figure 10. This is a positive result showing participants are able to determine complex shapes, without a priori knowledge of what to expect, in relatively short periods of time. Our analysis found no statistically significant difference between the time-to-completions for the two hinting modes $(T = 0.761 < T_{crit}^{\alpha=0.05} = 2.440)$, which corresponds with earlier results.

5.2 Indoor exploration application

In the first map, the participant interacted with the device for approximately 14 minutes before indicating he was confident about the layout of the space. The participant found all but one of the POI markers (the washroom), which he was able to locate after an additional 1:20 (one minute and twenty seconds) of interaction.

In the second, more complex map, the participant interacted with the device for approximately 10 minutes before indicating he was confident about the layout of the space. We attribute this increase in speed despite the more complex map to improved familiarity with the interface. The participant found all except the west-side break room and conference room. The participant took an additional 1:10 to find the conference room, and an additional 3:10 to find the break room. When asked for directions from the attorney's office to the break room, the participant correctly provided turn-by-turn directions and named the POI markers for the route via the "stairs down" path. The participant indicated that he suspected there could be a second path via passing the conference room, but was not confident.

Qualitatively, this experiment suggests that users are able to piece together multiple screenfuls of content to perceive

specific preferences for sonification modes under different environmental conditions (*e.g.*, noisy rooms or outdoors).

Our calculations also do not indicate statistically significant improvements in accuracy between blocks 1 and 2 for the line and area hinting sonification modes $(T = 0.458 < T_{crit} = 2.44)$. We suspect two contributing factors in this result. First, we believe that more time is needed for users to improve their skill with the interface. Some users commented that they believed more time and practice would improve their ability to use the interface. Second, from our observations, we speculate that some participants were mentally fatigued by the second block. Some of the participants were more senior in age, and some performed the experiment after work hours. These factors could impact the participants' concentration and accuracy. Future experiments could split the study into multiple sessions covering multiple days to reduce the impact

(a) line hinting sonification

(b) area hinting sonification

Figure 9: Summary of percentage answered correctly. Bars are means with 95% confidence interval

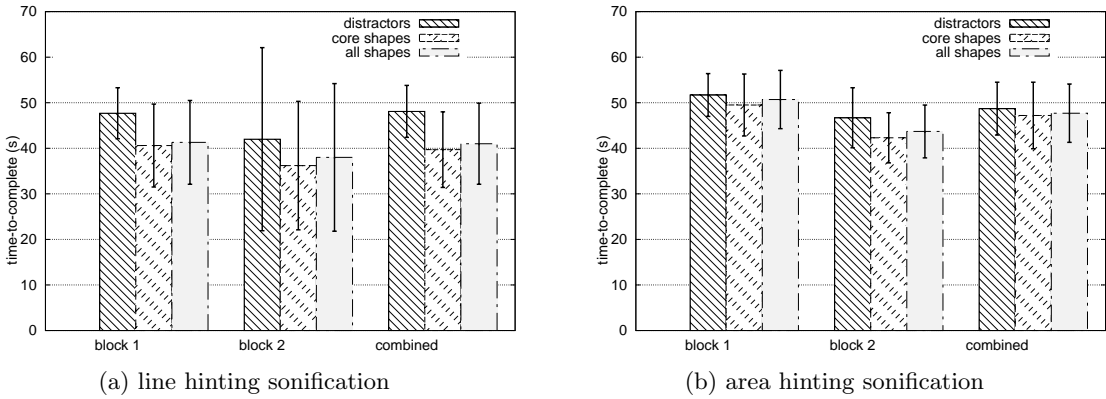(a) line hinting sonification

(b) area hinting sonification

Figure 10: Summary of time to answer correctly. Bars are means of median timing with 95% confidence interval

a much larger map. The participant noted that the larger map was intuitive to piece together because it utilized a similar skill to what is currently available in GPS navigation systems.

In post-experiment questioning, the participant was positive in his reaction to the application, stating that "if I were sitting in an airplane and had the time to look at the hotel lobby before I got there, I'd definitely do that." The participant went on to detail "this one particular hotel in Anaheim that I wish I could have had [this]. There was [sic] four different banks of elevators, and they were all in the middle of the lobby and the lobby was *huge*. It was all on angles; even sighted people had trouble."

With regards to the interface, the participant indicated that the largest difficulty he had was with the panning operation. Interestingly, the participant preferred to have the primary sensing finger stationary, and to use a secondary finger to slide the map under the primary finger. We suspect this preference may be because it is similar to the panning mode utilized by GPS navigation devices available today, where the map pans under a centrally anchored cursor.

6. DISCUSSION

In this section we present some observations of participant reaction and usage of the Timbremap interface. When comparing the two hinting modes, most participants remarked that the area hinting mode was much more aesthetically pleasing. Unfortunately, we believe this hurt the performance results for this interface due to the setting of our experiment. Our experimental test setting was "exam-like" in many ways, including timed progression and multiple-choice questions. Several users remarked that as they mentally fatigued, they often took time to enjoy the pleasing background sound effects in the area interface, despite the delay being detrimental to the task at hand.

Some participants noted that the area sounds were too different. We then tried changing them to be more subtly close (wind, beach waves, rain drops, hum of an air-conditioner). Other participants who tried the more similar area sounds remarked they would like the sounds to be more distinctly different. We suspect that the best sounds will depend on user context as well as individual user preferences.

Despite the line hinting mode being less aesthetically pleasing, many users preferred it over area hinting because the sonification provided explicit feedback for what to

	block1	block2	combined
distractor 1	45.8	40.5	45.3
distractor 2	46.4	38.0	43.8
distractor 3	46.0	51.5	50.3
coreshape 1	36.8	20.3	28.2
coreshape 2	37.1	34.8	36.6
coreshape 3	43.6	33.7	43.8
coreshape 4	49.2	43.8	46.2
coreshape 5	45.2	38.5	45.0
coreshape 6	45.5	47.5	48.3

(a) line hinting sonification

	block1	block2	combined
distractor 1	51.5	48.5	50.8
distractor 2	50.0	49.7	51.5
distractor 3	46.0	40.8	42.8
coreshape 1	38.2	35.0	37.8
coreshape 2	45.3	45.0	46.5
coreshape 3	46.0	46.0	47.8
coreshape 4	54.2	50.0	53.5
coreshape 5	53.5	52.5	50.0
coreshape 6	54.5	55.5	55.7

(b) area hinting sonification

Table 2: Median completion time distribution (on-time and correct answers).

do. Some participants remarked that this explicit hinting helped them stay focused. Removal of the diagonal hints initially used made the interface simpler and easier to learn. However, we observed that this directional restriction resulted in confusion in certain situations. For example, with some diagonal and curved paths, the hinted direction is not on the shortest perpendicular back to the path. We suspect this variability in finger distance to return to the path may result in user confusion.

We do not have sufficient data to conclude whether or not there is a correlation between braille reading ability and performance. However, based on anecdotal observation of the participants, there appears to be no correlation between braille reading ability and performance on the interface. We speculate that spacial positioning skills are likely to be a more important factor in determining whether a user performs well with the Timbremap interface.

There is little doubt that regardless of natural ability, the Timbremap interface is one that must be learned and practiced. Despite not seeing an appreciable improvement in interaction speed between the first and second blocks of each hinting mode, many participants speculated that additional training would likely improve their confidence, accuracy, and speed with the interface. We leave this extended study of improvement over training and time to future work.

7. CONCLUSIONS AND FUTURE WORK

This paper presented the motivation, design, and evaluation of an indoor map exploration sonification interface for visually impaired users on touch-screen mobile devices. This paper contributes an implementation and user study with visually impaired participants using two sonification hinting modes, both of which were effective in conveying shape geometry to participants. In addition, we qualitatively show that a participant is able to use the presented sonification interface for an indoor floor-plan

exploration application, and is able to piece together a map larger than can be displayed in one screenful.

For the sonification interface, we are planning to explore the effectiveness of the technique on medium and larger-sized touch interfaces such as tablets and table-tops. Specifically, whether larger spacial surfaces impact the speed and accuracy of users with this sonification interface. For the indoor floor-plan exploration application, the software can be extended to take advantage of built-in magnetometers to auto-rotate the map or indicate the user's facing direction. Further studies would be needed to explore techniques for conveying map rotation to the user.

One capability that the area hinting mode enabled is support for multi-touch feedback. Our prototype supported multi-touch feedback for multiple touch points. Due to time restrictions during the experiment, we opted not to introduce this additional complexity on our participants. However, we believe there are significant opportunities for exploring multi-touch feedback in both sonification modes in Timbremap, as well as exploring whether the addition of multi-touch provides a significant advantage for one mode over the other.

The current implementation of the Timbremap indoor exploration application loads floor-plans hand tailored for the evaluation. We envision users using high-speed cellular or WiFi to load detailed map and floor-plan information from the web on demand. These maps could be produced by a variety of means, such as image recognition of existing floor-plans, or manual annotation by volunteer efforts.

Further potential for future work includes developing a system which notifies visually impaired users as they approach hazards such as sidewalk construction detours. Timbremap could enable users to explore the characteristics of the detour and enable them to plan how best to traverse or avoid the hazard.

8. ACKNOWLEDGEMENTS

9. REFERENCES

[1] J. L. Alty and D. I. Rigas. Communicating graphical information to blind users using music: the role of context. In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 574–581, New York, NY, USA, 1998. ACM Press/Addison-Wesley Publishing Co.

[2] B. Blasch, W. Wiener, and R. Welsh. Foundations of orientation and mobility, second edition. In *American Foundation for the Blind*, pages 291–293, 1997.

[3] A. Crossan and S. Brewster. Multimodal trajectory playback for teaching shape information and trajectories to visually impaired computer users. *ACM Trans. Access. Comput.*, 1(2):1–34, 2008.

[4] T. Dingler, J. Lindsay, and B. N. Walker. Learnabiltiy of sound cues for environmental features: Auditory icons, earcons, spearcons, and speech. In *Proceedings of the 14th International Conference on Auditory Display*, Paris, France, 2008.

[5] A. Hub, J. Diepstraten, and T. Ertl. Design and development of an indoor navigation and object identification system for the blind. In *Assets '04:*

Proceedings of the 6th international ACM SIGACCESS conference on Computers and accessibility, pages 147–152, New York, NY, USA, 2004. ACM.

[6] D. Jacobson. Navigating maps with little or no sight: A novel audio-tactile approach. Technical report, 1998.

[7] S. K. Kane, J. P. Bigham, and J. O. Wobbrock. Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. In *Assets '08: Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*, pages 73–80, New York, NY, USA, 2008. ACM.

[8] F. C. Y. Li, D. Dearman, and K. N. Truong. Virtual shelves: interactions with orientation aware devices. In *UIST '09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pages 125–128, New York, NY, USA, 2009. ACM.

[9] http://www.loadstone-gps.com/.

[10] J. M. Loomis, R. G. Golledge, R. L. Klatzky, J. M. Speigle, and J. Tietz. Personal guidance system for the visually impaired. In *Assets '94: Proceedings of the first annual ACM conference on Assistive technologies*, pages 85–91, New York, NY, USA, 1994. ACM.

[11] J. R. Marston, J. M. Loomis, R. L. Klatzky, R. G. Golledge, and E. L. Smith. Evaluation of spatial displays for navigation without sight. 3(2):110–124, 2006.

[12] P. Parente and G. Bishop. Bats: The blind audio tactile mapping system. In *ACMSE*, 2003.

[13] D. A. Ross and B. B. Blasch. Wearable interfaces for orientation and wayfinding. In *Assets '00: Proceedings of the fourth international ACM conference on Assistive technologies*, pages 193–200, New York, NY, USA, 2000. ACM.

[14] B. N. Walker and J. Lindsay. The effect of a speech discrimination task on navigation in a virtual environment. In *Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society (HFES2006)*, pages 1538–1541, San Francisco, CA, 2006.

[15] B. N. Walker, A. Nance, and J. Lindsay. Spearcons: speech-based earcons improve navigation performance in auditory menus. pages 63–68, London, UK, 2006. Department of Computer Science, Queen Mary, University of London, UK, Department of Computer Science, Queen Mary, University of London, UK.

[16] J. Wilson, B. Walker, J. Lindsay, C. Cambias, and F. Dellaert. Swan: System for wearable audio navigation. In *Wearable Computers, 2007 11th IEEE International Symposium on*, pages 91–98, Oct. 2007.

[17] K. Yatani and K. N. Truong. Semfeel: a user interface with semantic tactile feedback for mobile touch-screen devices. In *UIST '09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pages 111–120, New York, NY, USA, 2009. ACM.

[18] S. Zhao, P. Dragicevic, M. Chignell, R. Balakrishnan, and P. Baudisch. Earpod: eyes-free menu selection using touch input and reactive audio feedback. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1395–1404, New York, NY, USA, 2007. ACM.