

# A Wearable Multimodal Interface for Exploring Urban Points of Interest

Antti Jylhä <sup>†</sup>, Yi-Ta Hsieh <sup>†</sup>, Valeria Orso <sup>‡</sup>, Salvatore Andolina <sup>†</sup>, Luciano Gamberini <sup>‡</sup>,  
Giulio Jacucci <sup>†§</sup>

<sup>†</sup>Helsinki Institute for Information Technology HIIT, Department of Computer Science, University of Helsinki, Finland  
first.last@helsinki.fi

<sup>§</sup>Helsinki Institute for Information Technology HIIT, Aalto University, Finland

<sup>‡</sup>Human Inspired Technology Research Centre HIT, Department of General Psychology, University of Padova, Italy  
first.last@unipd.it

## ABSTRACT

Locating points of interest (POIs) in cities is typically facilitated by visual aids such as paper maps, brochures, and mobile applications. However, these techniques require visual attention, which ideally should be on the surroundings. Non-visual techniques for navigating towards specific POIs typically lack support for free exploration of the city or more detailed guidance. To overcome these issues, we propose a multimodal, wearable system for alerting the user of nearby recommended POIs. The system, built around a tactile glove, provides audio-tactile cues when a new POI is in the vicinity, and more detailed information and guidance if the user expresses interest in this POI. We evaluated the system in a field study, comparing it to a visual baseline application. The encouraging results show that the glove-based system helps keep the attention on the surroundings and that its performance is on the same level as that of the baseline.

## Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces

## Keywords

Wearable computing; multimodal interaction; cultural heritage; tactile feedback; auditory icons

## 1. INTRODUCTION

Traditionally, exploration of points of interest (POIs) in cities has relied on printed materials such as maps and guide books. The smartphone era is effectively rendering these techniques obsolete, with a plethora of available interactive maps and tourist guide applications. Techniques range from

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).  
*ICMI 2015*, November 9–13, 2015, Seattle, WA, USA.

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

© 2015 ACM. ISBN 978-1-4503-3912-4/15/11 ...\$15.00.

DOI: <http://dx.doi.org/10.1145/2818346.2820763>.



Figure 1: Conceptual illustration of exploring urban points of interest. Instead of visual markers, our system uses audio-tactile feedback to notify the user of POIs in the vicinity and to guide the pointing hand of the user to locate the POIs.

visual augmented reality (e.g., [36], [29], [14]) to mixed reality approaches for displaying content on 2D maps (e.g., [23], [2]). More complex interaction techniques have also been developed for mobile 3D maps (e.g., [24], [25], [18]). In addition, location-aware mobile applications, such as Foursquare<sup>1</sup>, are redefining the way tourists explore new places.

However, hand-held visual interfaces may suffer from issues such as screen reflections interfering with content visibility [16, 37] and require visual attention on the device, which results in sharing cognitive resources between the device and the environment [6, 26]. This may disconnect the user from the surroundings and hamper the user experience [3]. A remedy for these issues can be provided by head-mounted displays or non-visual interfaces.

Several approaches have been proposed making use of the auditory and/or haptic modalities to enhance the user experience, to enable the use of the system in situations when visual attention cannot be on the device, and to cater also for visual disabilities. However, the vast majority of these systems have been designed for navigating a defined route or finding a specific place, thus limiting their use for truly exploring the city, which is also a relevant mode of sightseeing [5, 12]. This leads to our primary research question: how to design a non-visual interaction technique that supports exploratory, serendipitous discovery of POIs (see Fig. 1 for a conceptual illustration of the setting.).

<sup>1</sup><https://foursquare.com>

We present a novel location-aware, wearable audio-tactile system with an exploratory approach to find POIs. The interaction technique is designed for free exploration, without the need for pre-defined itineraries or direct contact with a mobile device. The system, based on a sensor-equipped tactile glove, gives audio-tactile cues indicating the *category* of a POI when the user is approaching it. If the user shows explicit interest in the POI, the system can guide the attention of the user towards the POI by audio-tactile feedback. Tour guide-like spoken descriptions give more information. Our main hypothesis, which is assessed through a comparison study, is that the user experience of the proposed system outperforms a visual mobile application. Based on the findings, we discuss the implications for the design of non-visual POI exploration. We contribute to the field of multimodal interaction with a novel non-visual technique for exploring POIs in cities, its evaluation in a field study, and implications for the design of non-visual tour guides.

## 2. BACKGROUND

In this section, we cover the related work on mobile and wearable tactile and auditory guidance systems.

### 2.1 Vibrotactile Guidance

Deploying vibrotactile feedback for guidance has been studied with both mobile devices and wearables. In outdoor contexts, the predominant use cases have been waypoint navigation and target finding, for which proposed solutions include, e.g., belts [7] [34] [4], shoes [8], and mobile applications such as PocketNavigator [27], Tactile Compass [28], NaviRadar [31], and the Time Machine [33, 21]. These examples, however, tend to assume that the user is navigating towards a defined destination and lack support for exploration. Mobile device-based target scanning techniques by Ahmaniemi and Lantz [1] and Robinson et al. [30] differ in the sense that they do not guide the user towards a specific destination, but instead enable scanning targets by pointing with a mobile device. The laboratory-based findings of the latter study [30] suggest that while people tend to slightly slow down their walking pace when looking for targets based on haptic feedback, they are still able to find the targets while maintaining awareness of their surroundings.

Other form factors of wearable tactile guidance for pointing, such as gloves, have mainly been applied in indoor settings (e.g., [32], [19]), in which hand movements can be tracked visually, e.g., with depth cameras. While these feedback techniques can be efficient in guiding the pointing hand, the requirement for visual tracking fits poorly in outdoor contexts. This can be overcome by integrated sensors [13] that track the hand orientation.

Hornecker et al. [12] presented a serendipitous tour guide mobile application making use of vibrotactile cues. In their approach, the system is initialized by the places and attractions the user wishes to see. They utilize proximity model-based vibrotactile cues, when the user is approaching a relevant place. They also experimented with generic notification sounds, but found them not to be useful in pilot tests.

Our design makes use of both proximity-based vibrotactile notifications and more refined vibrotactile guidance for pointing at the recommended target. Our technique does not assume the user has any pre-defined destination, enabling true exploration of the city, while still being able to provide fine-grained guidance on demand to a specific POI.

### 2.2 Auditory Guidance and Recommendations

In addition to haptics, audio has been used in many ways in tour guide applications and related systems. The use of sound in these systems can be categorized to guidance cues [11, 20, 15], notifications [15], contextual sounds [22, 33, 21], and spoken information content [21].

Directional guidance sounds are most often seen in navigation systems. Among the first was AudioGPS [11], which utilized spatial panning and Geiger counter metaphor for guiding the navigation. A related approach was deployed by Kostiaainen et al. [15] in an audio-based mobile journey planner. The application used stereo panning cues to guide the person on the move, e.g., when riding a bicycle, to inform whether they deviated from the correct route. The application also made use of various notification sounds, e.g., to indicate that the person should exit the bus at the next stop. A different guidance scheme was used by Liljedahl and Lindberg [20], who developed a target scanning application with auditory proximity cues to relevant places in the direction where the mobile device was pointing.

The Lund Time Machine [33, 21] mobile application makes use of ambient sounds related to the history of a location to contextualize the user to the setting. The application uses by default vibration bursts for feedback when scanning the environment with the device, but also offers the possibility to use sound or speech for guidance.

As an example of a technique for serendipitous discovery, McGookin et al. [22] developed Audio Bubbles, an application that uses virtual geolocated spheres as auditory "homing beacons". When a person enters an audio bubble, they start hearing a sound that relates to the POI in the center of the bubble. The approach empowers the traveler to choose which audio bubbles they want to attend to. In a somewhat similar fashion, Vazquez-Alvarez et al. [35] studied the use of earcons and spatial audio in an augmented reality "sound garden." The system was tuned for presenting information on multiple targets at the same time, making use of source separation enabled by the spatial audio. Their results indicate that two simultaneous sound sources are distinguishable, but with three or more the task becomes difficult.

Ankolekar et al. [3] compared different cues for serendipitous discovery of POIs. The cues were musicons (music snippets related to a POI), auditory icons (sound samples related to the category of the POI [9]), speech (name and address of the POI), visual (name, address, category, and picture of the POI), and non-visual cues combined. The results indicate that ambiguous cues such as musicons and auditory icons result in a weaker confidence in identifying the POI, but give a greater sense of autonomy to the user. Speech cues were the fastest for identifying the POIs, while auditory icons were deemed more challenging to identify.

In summary, previous solutions for tactile or auditory guidance in cities typically do not support free exploration and serendipitous discovery of POIs. While a few counter-examples exist, they tend to either lack more fine-grained guidance, do not provide more detailed information on the POI, or present the user with potentially too many alternatives at a time. In addition, many systems require carrying a mobile device in hand and rely on visual feedback. Our system not only supports serendipity in discovering the POIs but also emphasizes users' autonomy to choose which recommended POIs they are interested in.

### 3. DESIGN

Our system supports exploring POIs in cities through two functional steps in the interaction technique:

1. Serendipitous discovery of POIs is supported by location-aware, contextual audio-tactile cues, which deliver the information of the category of a nearby POI. The user has the option to skip or pursue the recommended POI.
2. Locating the POI is supported by audio-tactile guidance that guides user’s pointing hand to point at the direction of the POI in question.

The interaction technique has been designed to be paired with a backend, which takes into account the preferences and POI history of the user to filter and personalize the content. In this paper we focus on the interaction technique.

#### 3.1 Tactile Glove

The system is based on a tactile glove (see Fig. 2), which is equipped with vibrotactile actuators and a set of sensors, and an Android mobile device. The glove has been constructed of thin elastic fabric — to make it adaptable to different hand sizes — and several electronic components. Two Arduino microcontrollers (Arduino Pro Mini) process the sensor signals and commands, and control the vibrotactile feedback. For sensing hand orientation, the glove is equipped with a 9-axis Inertial Measurement Unit (IMU, InvenSense MPU-9150), which consists of a gyroscope, an accelerometer, and a compass. Flexible bend sensors (Spectra Symbol flex sensor) are deployed on three fingers (thumb, index, and middle finger). A set of three vibrotactile actuators (Precision Microdrives 10mm shaftless vibration motor) is mounted for providing notification cues and directional guidance on the thumb, index, and middle fingers. The glove communicates wirelessly over Bluetooth with an Android mobile phone running the application software. The glove has been built for right-handed interactions.

A glove has been chosen as the form factor of our system for several reasons. Firstly, a glove allows natural hand gestures, which enables an embodied experience in interacting with the environment. Point-by-finger interaction naturally fits our tourism context, and information associated with a specific POI could be accessed via hand gestures projected on the POI itself. Moreover, fingers are more sensitive for tactile stimuli than, for example, the forearm [17].

#### 3.2 Proximity-Based Cues and Contextual Information

The core idea of the interaction technique is to allow the user to freely walk around in the city and discover POIs. This is achieved by proximity-based audio-tactile notifications. The user’s location is tracked by the GPS sensor on the mobile device. Each POI is associated with a proximity radius which is determined by its physical size and visibility. For example, a prominent palace has a larger radius than an ordinary shop. Considering using the real environment as the genuine visual cue of our non-visual system, the visibility of the POI also affects its proximity radius. If the POI is only visible after passing a corner, the radius is reduced to the distance between the corner and the POI. When the user crosses a proximity radius of a specific POI, the actuators on the glove vibrate to notify that a POI is nearby. Simultaneously, a contextual auditory icon is played through headphones to indicate the category of this POI. In the current



Figure 2: The tactile glove showing both sides.

Table 1: Description of the auditory icons designed for the system. The icons were formed as overlaid collages of multiple archetypal sound recordings related to the category.

POI Category	Sound Material
Architecture	Hammering, sawing
Shop	Old cashier machine, coins dropped on metal tray
Cafe	Espresso machine steam, grinding beans, pouring liquid, spoon hitting a coffee cup, sipping coffee
Restaurant	Knives and forks hitting each other and a plate
Religious	Church bells, Gregorian chanting

setup, we have defined five categories of POIs: architecture, shop, cafe, restaurant, and religious. An archetypal auditory icon was designed for each category. The sound material for the icons is summarized in Table 1, and was carefully chosen to yield designs that are easily distinguishable from each other and from typical environmental sounds. The duration of each auditory icon is approximately 2.5 seconds.

The audio-tactile notification lasts for 15 seconds, with the auditory icons playing back repeatedly, accompanied by vibration bursts (100 ms activation with 400 ms silence) on the glove. During this period, the user can make a selection gesture by bending the index finger. Alternatively if they are not interested, they can reject the recommendation by bending the thumb or ignore the cue until it fades out. Selection gesture results in hearing the spoken name of the POI, followed by a brief (max 10 seconds) spoken description of the POI. This spoken introduction is meant for giving the user more autonomy over choosing interesting POIs, and they can skip the POI if they find it uninteresting.

#### 3.3 Guidance for Pointing at the Target

When the user has listened to the initial description, they start getting directional guidance to spot the POI. The assumption is that the POI is in view of the user. The rationale behind this step is that there may be several POIs of one category in close proximity to each other. For example, it is common to have multiple restaurants or cafes side by side, and finding the correct one may require effort.

The directional guidance on the glove assumes that the user starts pointing with their glove-wearing hand with the palm facing down, browsing the environment to find the POI. The pointing interaction was chosen, because pointing at landmarks is a common phenomenon among tourists. Based on the sensor readings of the IMU on the glove and the GPS locations of both the user and the POI, the system computes the angular difference between the pointing direction and the actual direction of the POI. This information is then applied to the directional audio-tactile cues. When the user is very close to the POI, poor GPS accuracy might result in inaccurate guidance. To compensate for potential locally sub-optimal accuracy of the GPS, if the user is estimated to be within 10 m of the POI, the system stops giving guidance and simply plays back a spoken audio piece: "You are now in front of" followed by the name of the POI.

The vibrotactile guidance follows the metaphor of pulling the hand towards the vibrations. If the user is pointing too much to the right, they will feel continuous vibrotactile bursts (100 ms activation with 400 ms silence) on their thumb, indicating they need to point more to the left. Conversely, the guidance cue is presented on the middle finger to guide pointing to the right. When the user is pointing at the target, all three actuators vibrate with a faster rate (100 ms activation and 200 ms silence). The spatial left/right mapping for the tactile cue was chosen because it gives an immediate indication of the direction, which would be lacking in an intensity-based hot/cold metaphor approach.

The angular width of the target ranges from 23 to 45 degrees depending on the physical size of the POI, the current distance, and the run-time GPS accuracy. The minimal angular width was derived from a pilot study, in which users repeatedly pointed with the glove at four marked dots in different directions. The sensor readings were sampled, and the mean deviation was set as the minimal target size.

The audio guidance works together with the vibrations to enhance the perception of the directional cue. An amplitude- and phase-modulated wideband tone, with the fundamental frequency of approximately 200 Hz, is played through the headphones with stereo panning, guiding pointing to left and right with the same metaphor as with the vibrations. The pulsation frequency of the modulated sound is much faster than that of the vibration bursts, in order to minimize possible negative cross-modal rhythm perception effects that could result if the two cues have a similar pulsation tempo but are slightly out of sync. When the user is pointing at the correct target, they hear the spoken name of that target and feel the vibrotactile on-target cue.

Once the user is pointing at the correct target and makes a selection gesture, they hear a longer spoken description of the POI. This description may include, e.g., facts related to the history of a building, or the cuisine of a restaurant. The user can continue the exploration when the spoken description is finished, or make a rejection gesture to terminate.

## 4. EVALUATION

We evaluated the interaction technique in a guided, task-based experiment. In this section, we present the venue, method, procedure, and materials used in the evaluation.

### 4.1 Study Setting

The experiment was conducted in the center of a mid-sized city, in a touristic area with a lot of applicable POIs. We

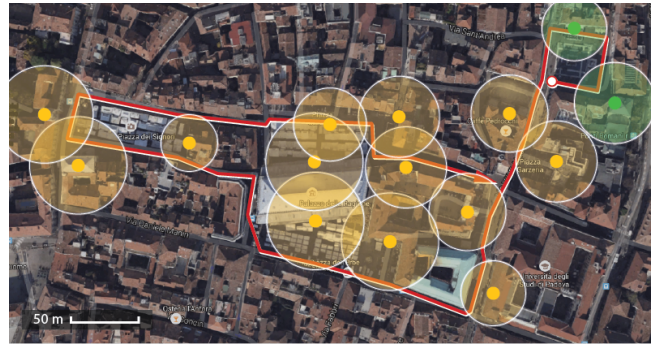


Figure 3: A map showing the locations of POIs and their associated proximity radius. The red line indicates the route users took in the study. The two targets in the top right corner were used for training.

chose 13 POIs from an area of roughly 0.12 km sq, and two extra POIs for training purposes outside the main area (See Fig.3). The POIs included, e.g., a church, two cafes, two shops, a statue, and facades and towers of old buildings.

For each POI, we prepared tour guide-like content related to the POI in both written and spoken format. The content included the name of the POI, a short introduction (roughly 10 seconds) to the POI, and a longer (approximately 30 seconds) description. Each content was also associated with the corresponding auditory icon representing the category of the POI. To ensure that all the contents were equally accessible to all participants and to avoid them going outside the GPS coverage, we defined a route to be followed within the area.

## 4.2 Experiment Design

The evaluation followed between-subjects design with two conditions. The primary condition involved using the proposed glove-based interaction technique for exploring the POIs. As a baseline condition, we implemented a mobile application manifesting relevant features competitive with commercial products available on the latest market. Figure 4 depicts users in both conditions.

### 4.2.1 Baseline Application

The baseline (Fig. 5) mimics the behavior of popular mobile apps for city exploration. Two main features presented here are push notification of recommended POI, as in Four-square<sup>2</sup>, and showing direction with compass, as in TripAdvisor City Guides<sup>3</sup>.

When a new POI is nearby, the app triggers a vibration and shows a notification suggesting to check out the POI (Fig. 5a). Users can decide to either discard the notification or to see more info about the POI (Fig. 5b), including picture, textual, and audio description. Similarly to our main system, the baseline also provides directional guidance to spot the POI, but using visual feedback as opposed to tactile feedback with an arrow pointing towards the POI (Fig. 5c).

## 4.3 Participants

Overall twelve people (five women) volunteered for the experimental trials, half of whom were assigned to the baseline group (three women), whilst the other half (two women)

<sup>2</sup><https://foursquare.com>

<sup>3</sup><http://www.tripadvisor.com/apps>



Figure 4: Participants operating the baseline application (left) and the tactile glove (right).

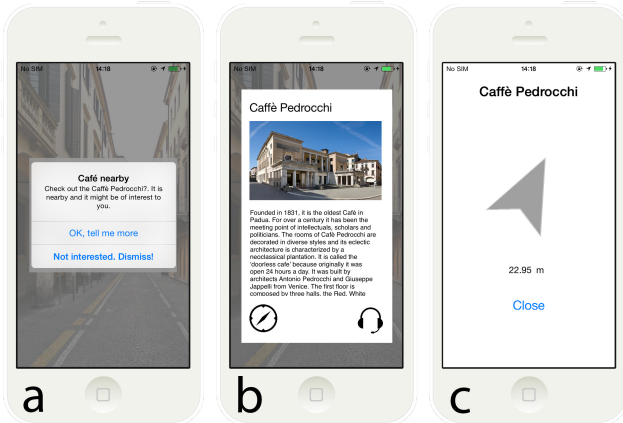


Figure 5: Baseline. a) A new POI nearby triggers visual and tactile (vibration) notification. b) Main view with image and description of the poi; directions and audio description are accessible from here. c) Directions View

used the tactile glove. They were all right-handed and had normal or corrected-to-normal vision. All participants were tourists recruited in the field and all of them reported to have none or very limited knowledge of the city. The mean age of the total sample was 34 years ( $SD = 17.9$ ); the baseline group had an average age of 34.66 years ( $SD = 17.97$ ) and the experimental group 30.66 years ( $SD = 19.49$ ).

#### 4.4 Apparatus

The participants of the proposed interaction technique wore the tactile glove in their right hand. Participants carried a neck pouch with an Android device (Samsung Galaxy S3) inside. The phone was used to run an application with the experiment logic, communicating wirelessly with the glove. Audio feedback was presented through an over-ear headset connected to the phone. Although an in-ear headset would have ensured better acoustic insulation, an over-ear headset was preferred due to hygienic reasons and to not disconnect the subject from the environment.

The participants of the baseline condition used an iPhone 5, with the baseline application pre-installed, and the same headphones as the other participants.

#### 4.5 Procedure

Participants first gave informed consent to participate in the study and then filled in a short background questionnaire. Next, the experimenter explained how the device,

either the tactile glove or the baseline application, depending on the experimental assignment, worked. When the user confirmed that the system functioning was clear, she received the device and the training phase began. This stage had the aim of familiarizing the user with the device and the interaction modalities involved. The training consisted of a guided exploration of a small square, on which there were two POIs that the participant had to explore. Participants were free to repeat the training phase if needed. When the user reported as ready, the experimental trial started. During the trial, the participant was instructed to follow the experimenter along the pre-defined route (approximately 800 meters in length). During the experimental session, the participant was asked to avoid any interaction with the experimenter, with the exception of reporting which building or monument they perceived to be attending. Participants using the iPhone were free to hold it in their hand or keep it in their pocket. When the participant reached the end of the route, she was asked to fill in a short questionnaire and to answer a brief semi-structured interview.

#### 4.6 Data Collection

The measures recorded during the test include observations, questionnaires, and logged data. The entire trial was video-recorded to allow off-line analysis of the users' behavior. Furthermore, during the testing session, the correct identification of each POI was recorded online by the experimenter, who explicitly asked the participant to indicate the location to which the guidance referred.

In addition to a post-test interview, two questionnaires were administered. One was filled before the experimental session and aimed at collecting background information regarding participants' age, origin, level of knowledge of the city, and how participants were used to sightseeing a city as tourists. The questionnaire completed after the trial explored users' impressions of the user experience. The survey explored participants' general feeling of naturalness and pleasantness in using the devices (3 items) and their impressions regarding the feedback (5 items). Finally 3 items explored the fear of appearing strange in the eyes of other people while using the device. Participants using the tactile glove were asked to complete a longer version of the questionnaire, in which they were also asked to evaluate the audio signals they received (5 items) and the physical sensations they had while wearing the tactile glove (5 items). Participants were asked to mark their level of agreement with each statement on a 5-point Likert scale. The final phase consisted of a brief semi-structured interview, in which participants were asked about their general impression of their use of the devices, what they found easy or difficult, what they liked, what they thought of the audio and vibrotactile signals and if they felt confident of the indications they received.

Both systems logged timestamped data that contained users' behavior including time spent in each stage, decisions of accepting/rejecting the notification, and distance to the POI, as well as environmental data like GPS accuracy.

### 5. RESULTS

In the evaluation we focused on the validation of the interaction technique and user experience. In terms of efficiency, the result shows that our non-visual interface performs as well as its visual counterpart. Positive user experience was

Table 2: Time spent on different stages in the interaction. Response time: the duration between getting and accepting the recommendation. Finding time: the time spent on finding the POI with guidance. Mean time in bold and standard deviation in parentheses.

Stage	Glove	Baseline
Response Time (s)	<b>5.9</b> (4.23)	<b>3.58</b> (2.04)
Finding Time (s)	<b>11.1</b> (14.25)	<b>13.48</b> (20.66)

achieved with both conditions. However, glove users were able to focus their attention on the environment, while baseline users paid much less attention on their surroundings. In this section, we present the results in detail.

### 5.1 Reacting to the Cues

Users’ interactions with the devices were logged by the system. The time required to respond to the notification and to find the POIs are reported in Table 2. In general, the response time of the glove condition is longer than the baseline, yet the difference is not significant according to a Mann-Whitney test ( $U=8.00$   $p=.109$ ). This can be explained, at last in part, by the fact that all participants in the baseline condition chose to hold the iPhone in their hand. Apart from the vibrotactile cue, which was the same in both conditions, the baseline application showed immediately explicit information regarding the POI on the screen, while the glove interface required the user to listen to the playback of auditory icons before making a selection and accessing introductory information of the POI. On-site observation also contributed to explain the difference. While glove users tended to slow down smoothly after the notification, baseline users tended to abruptly stop to interact with the phone. Regarding the time needed to find a POI, users in the glove condition were faster compared to those in the baseline group. Again, the difference was not significant according to a Mann-Whitney test ( $U=12$   $p=.96$ ).

Since participants were free to accept or reject the notification, four glove users and two baseline users actively rejected or passively ignored some of the notifications. We also observed that not all POIs were successfully recommended to users due to poor GPS coverage in some conditions. The recorded accuracy of GPS was typically good, ranging from 5 m to 30 m with an average of 12.37 m. With a single user at two POIs the accuracy dropped to 50 m and 100 m.

### 5.2 Identifying the POIs

An index of POI identification accuracy was computed for each participant, consisting of the percentage of the number of correct POI identifications with respect to the total number of POIs delivered by the system. Given the limited sample size, a Mann-Whitney test was run to compare the POI identification accuracy in the two conditions. The analysis revealed no significant difference between the conditions ( $U=18.00$   $p=1$ ), with an average success rate of 89.24% ( $SD=10.6$ ) for the group exploring the city center with the baseline application and an average success rate of 88.35% ( $SD=11.62$ ) for the group wearing the tactile glove. This finding suggests that the tactile glove is able to guide the user to locate a POI with the same accuracy as a more familiar application that provides an extra visual clue, a picture, for identifying the target.

## 5.3 General User Experience

Referring to the opinions that emerged from the Likert-scale questionnaires, no significant difference in user experience was found. In particular, both groups showed a positive attitude toward the applications: both the glove and the baseline were considered unobtrusive (baseline  $M=2$ ; glove  $M=1.6$ ), natural (baseline  $M=3.83$ ; glove  $M=3.83$ ), pleasant to use (baseline  $M=4$ ; glove  $M=4.16$ ), and not interfering with the main task (baseline  $M=2.16$ ; glove  $M=1.6$ ). Vibration signals were well received in both conditions, being evaluated as attention getting (baseline  $M=4$ ; glove  $M=4.16$ ), accurate (baseline  $M=3.83$ ; glove  $M=4$ ), and easy to interpret (baseline  $M=4.16$ ; glove  $M=4$ ). Furthermore, participants considered them pleasant (baseline  $M=4$ ; glove  $M=3.6$ ) and not disturbing (baseline  $M=1.16$ ; glove  $M=1.3$ ). Interestingly, participants in both conditions showed only low concern of looking strange to other people (baseline  $M=2.16$ ; glove  $M=2.6$ ), even if the proposed technique is all but common. It is worth noticing that participants of both groups did not feel hampered by the device they were employing to explore the city (baseline  $M=2.16$ ; glove  $M=2.3$ ).

All the participants appreciated the intuitiveness and ease of use of both the tactile glove and the baseline application. Some participants in the tactile glove condition commented spontaneously after the trial “I enjoyed it!” One user commented “I liked it! It really works, it was very precise.” Regarding the feeling of wearing the glove, users found the glove light-weight ( $M=4.66$ ,  $SD=0.51$ ), comfortable ( $M=4$ ,  $SD=0.63$ ), and well-suited to their hand ( $M=4.33$ ,  $SD=0.51$ ). Finally, the tactile glove was not found bulky ( $M=2$ ,  $SD=0.63$ ) or rigid ( $M=1.83$ ,  $SD=0.4$ ), nor to limit the movements of the hand ( $M=2.16$ ,  $SD=0.75$ ).

### 5.4 Attention on the Surroundings

In general, participants praised the alerts coming only in the proximity of a certain POI, instead of having to pay attention to a paper map or a guide during the entire walk. Two participants in the baseline group explicitly said they felt detached from the environment, as they were focusing their attention on the application: “I was certainly thinking of the smartphone and not looking so much at the city.” On the other hand, in the group of the tactile glove users, the feeling of being disengaged from the surroundings occurred only for one respondent, with one user reporting an opposite feeling “I felt like I was in a museum.”

When the exploration was led by the baseline application, participants seemed to divide their attention between the surroundings and the screen of the smartphone. Based on the video analysis, from the notification until finding the POI, they spent on average 70.27% ( $SD=13.25$ ) of the time looking at the screen (including, of course, the time to read the textual content). Participants tended to periodically check the screen of the phone while walking from a certain POI to the following one. When they received the notification of a POI in the proximity, they tended to stop abruptly and to focus their gaze on the screen. Even if they explicitly followed the directional cues provided by the baseline application to identify the POI, they usually then checked by comparing what they were seeing in front of them and the picture displayed on the screen. The screen seemed to grab participants’ gaze even when they were listening to the audio description.

Video analysis shows that glove users seemed to be having a walk around the city. When they got a notification of approaching a POI, they slowed down smoothly and simply concentrated on the vibrations. When the audio description was playing, they looked at the POI or took quick glances around. Interestingly, even when they showed some uncertainties in interpreting the tactile guidance, none of the participants ever looked down at the glove. Overall, the glove did not appear as an external device put onto the user's body, but rather seemed well integrated with the flow of actions. Users operated it with the same naturalness as if they were making a movement with their hand.

## 6. DISCUSSION

In previous studies, non-visual interfaces have not been considered as efficient as their visual counterparts (e.g., [3]). Our evaluation surprisingly indicates that the performance of the presented non-visual solution is not significantly different from the visual baseline application for locating the POIs. This might be explained by two considerations.

First, the users of the baseline application need to divide their attention between the visual guidance and the real surroundings, repeatedly switching their gaze between the screen and the environment, while the glove-based solution gives people the chance to keep their visual attention on the surroundings. Potentially, extending the arm to a "pointing posture" while holding the phone would better align the gaze, the screen, and the target, but this does not seem like an intuitive way of interacting with the smartphone. None of the participants applied this strategy.

Second, while there is no visual feedback from the glove-based system, the surroundings themselves can provide strong visual cues to accompany the audio-tactile guidance. The combined information of the auditory icon, the spoken name of the target, and the short spoken description provide the user with relevant information to be able to distinguish the POI visually from others. When facilitated by the audio-tactile feedback and the possibility of not having to look at the device, the proposed solution is actually slightly faster in the guidance step than the visual baseline.

However, efficiency is not always the highest priority [3]. In our case, achieving a more immersed experience where users' attention is on the real surroundings is a key consideration. In addition to the non-visual directional guidance, this immersion is in part also supported by the auditory icons, which augment the surroundings with contextual information. It should be noted that interpreting the auditory icons might not require much cognitive load but a certain time duration is necessary to be spent on listening to decode the information.

Regarding user experience, there were no significant differences in questionnaire responses. The lack of differences is actually encouraging, for example, in terms of social acceptability of operating the glove in the wild, since it was not considered significantly more obtrusive than operating the smartphone.

The glove proved to be a good form factor for "browsing" the city on the go. While other wearables such as shoes [8] or belts [7] could be used for guidance as well, the glove provides a nice integrated package for versatile interactions both as an input device using onboard sensors and as an output device via the tactile feedback. Furthermore, these

interactions make use of a standard practice in tourism, i.e., pointing at things in the surroundings.

## 6.1 Limitations

In this study, we used GPS for positioning. While it performs generally well in open areas outdoors, the positioning may be inaccurate in certain places such as in very narrow streets. For these situations, the positioning accuracy should be taken into account when notifying the user about new content in proximity. Clearly, if the interaction technique would be adapted for indoor use, for example in museums, a different positioning technique would be required, involving perhaps the reading of fiducial markers ([10]) or other specific approach tuned for those situations.

Since the non-visual interface provides no visual reference, in some situations the user might be unsure whether she is pointing at the correct POI. On a related note, in this study the system relied on manually determining the visibility of the POIs along the given path. We are currently working on determining POI visibility on the system side.

We did not encode any directional information in the auditory content, unlike some previous systems (e.g., [35]). We did not want to impose multiple directional guidance schemes on the user at this stage, but are considering pursuing this in the future. Another essential future development will be to integrate the interaction technique with a recommendation engine backend to personalize the content. Now that we have confirmed that the interaction technique itself works, in future studies we will work towards an integrated, multimodal system for exploring new places.

## 7. CONCLUSION

We proposed a novel interaction technique for exploring POIs in cities. The technique, based on audio-tactile feedback and a tactile glove, was evaluated in a field study and was shown to be as successful as a visual baseline application for locating and identifying the POIs. Interestingly, there were no major differences in questionnaire-based user experience between the two techniques, but observations and interviews revealed that the proposed technique let the user focus more attention on the surroundings than on the device compared to the baseline. Also, while the glove represents an unfamiliar technique and relies on pointing with the hand to operate, the results indicate that it was not considered to be more obtrusive to use than the baseline application, suggesting that its social acceptability does not pose significant problems compared to more standard mobile applications. In the future, we will integrate the interaction technique with a backend, which will personalize the POI content according to user preferences.

## 8. ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n°601139.

## 9. REFERENCES

- [1] T. T. Ahmaniemi and V. T. Lantz. Augmented reality target finding based on tactile cues. In *Proc. ICMI-MLMI*, pages 335–342, 2009.
- [2] S. Andolina, D. Pirrone, G. Russo, S. Sorce, and A. Gentile. Exploitation of mobile access to

- context-based information in cultural heritage fruition. In *Proc. BWCCA '12*, pages 322–328, 2012.
- [3] A. Ankolekar, T. Sandholm, and L. Yu. Play it by ear: a case for serendipitous discovery of places with musicons. In *Proc. CHI*, pages 2959–2968, 2013.
- [4] A. Asif, W. Heuten, and S. Boll. Exploring distance encodings with a tactile display to convey turn by turn information in automobiles. In *Proc. NordiCHI*, pages 32–41, 2010.
- [5] B. Brown and M. Chalmers. Tourism and mobile technology. In *ECSCW 2003*, pages 335–354, 2003.
- [6] L. Chittaro. Distinctive aspects of mobile interaction and their implications for the design of multimodal interfaces. *J. Multimodal User Interfaces*, 3(3):157–165, 2010.
- [7] J. B. F. V. Erp, H. A. H. C. V. Veen, C. Jansen, and T. Dobbins. Waypoint navigation with a vibrotactile waist belt. *ACM Trans. Appl. Percept.*, 2(2):106–117, Apr. 2005.
- [8] M. Frey. CabBoots: Shoes with Integrated Guidance System. In *Proc. TEI '07*, pages 245–246, 2007.
- [9] W. W. Gaver. Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*, 2(2):167–177, 1986.
- [10] A. Gentile, S. Andolina, A. Massara, D. Pirrone, G. Russo, A. Santangelo, E. Trumello, and S. Sorce. A multichannel information system to build and deliver rich user-experiences in exhibits and museums. In *Proc. BWCCA '11*, pages 57–64, 2011.
- [11] S. Holland, D. R. Morse, and H. Gedenryd. AudioGPS: Spatial audio navigation with a minimal attention interface. *Personal and Ubiquit. Comput.*, 6(4):253–259, 2002.
- [12] E. Hornecker, S. Swindells, and M. Dunlop. A mobile guide for serendipitous exploration of cities. In *Proc. MobileHCI*, pages 557–562, 2011.
- [13] Y.-T. Hsieh, A. Jylhä, and G. Jacucci. Pointing and selecting with tactile glove in 3d environment. In *Proc. Symbiotic 2014*, pages 133–137, 2014.
- [14] H. Jo, S. Hwang, H. Park, and J. Hee Ryu. Aroundplot: Focus+Context Interface for Off-Screen Objects in 3D Environments. *Computers & Graphics*, 35(4):841–853, 2011.
- [15] J. Kostiaainen, C. Erkut, and F. B. Piella. Design of an audio-based mobile journey planner application. In *Proc. Academic MindTrek Conf.*, pages 107–113, 2011.
- [16] E. Kruijff, J. E. Swan II, and S. Feiner. Perceptual issues in augmented reality revisited. In *Proc. ISMAR*, pages 3–12, 2010.
- [17] S. J. Lederman. Skin and touch. *Encyclopedia of Human Biology*, 7:51–63, 1991.
- [18] V. Lehtinen, A. Nurminen, and A. Oulasvirta. Integrating spatial sensing to an interactive mobile 3D map. In *Proc. 3DUI*, pages 11–14, March 2012.
- [19] V. Lehtinen, A. Oulasvirta, A. Salovaara, and P. Nurmi. Dynamic tactile guidance for visual search tasks. In *Proc. UIST*, pages 445–452, 2012.
- [20] M. Liljedahl and S. Lindberg. Sound parameters for expressing geographic distance in a mobile navigation application. In *Proc. Audio Mostly*, pages 1–7, 2011.
- [21] C. Magnusson, K. Rassmus-Gröhn, and D. Szymczak. Exploring history: a mobile inclusive virtual tourist guide. In *Proc. NordiCHI*, pages 69–78, 2014.
- [22] D. Mcgookin, S. Brewster, and P. Priego. Audio bubbles: Employing non-speech audio to support tourist wayfinding. In *Proc. HAID*, pages 41–50, 2009.
- [23] A. Mulloni, A. Dünser, and D. Schmalstieg. Zooming interfaces for augmented reality browsers. In *Proc. MobileHCI*, pages 161–170, 2010.
- [24] A. Nurminen. m-LOMA - a mobile 3D city map. In *Proc. Web3D '06*, pages 7–18, 2006.
- [25] A. Oulasvirta, S. Estlander, and A. Nurminen. Embodied interaction with a 3D versus 2D mobile map. *Personal and Ubiquit. Comput.*, 13(4):303–320, 2008.
- [26] A. Oulasvirta, S. Tamminen, V. Roto, and J. Kuorelahti. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. In *Proc. CHI*, pages 919–928, 2005.
- [27] M. Pielot, B. Poppinga, and S. Boll. PocketNavigator: vibro-tactile waypoint navigation for everyday mobile devices. In *Proc. MobileHCI*, pages 423–426, 2010.
- [28] M. Pielot, B. Poppinga, W. Heuten, and S. Boll. A tactile compass for eyes-free pedestrian navigation. In P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, and M. Winckler, editors, *Proc. INTERACT*, volume 6947 of *Lecture Notes in Computer Science*, pages 640–656, 2011.
- [29] M. Rantanen, A. Oulasvirta, J. Blom, S. Tiitta, and M. Mäntylä. InfoRadar: Group and Public Messaging in the Mobile Context. In *Proc. NordiCHI*, pages 131–140, 2004.
- [30] S. Robinson, P. Eslambolchilar, and M. Jones. Evaluating haptics for information discovery while walking. In *Proc. British HCI*, pages 93–102, 2009.
- [31] S. Rümelin, E. Rukzio, and R. Hardy. NaviRadar: a novel tactile information display for pedestrian navigation. In *Proc. UIST*, pages 293–302, 2011.
- [32] R. Scheibe, M. Moehring, and B. Froehlich. Tactile feedback at the finger tips for improved direct interaction in immersive environments. In *Proc. 3DUI*, pages 1–7, 2007.
- [33] D. Szymczak, K. Rassmus-Gröhn, C. Magnusson, and P.-O. Hedvall. A real-world study of an audio-tactile tourist guide. In *Proc. MobileHCI*, pages 335–344, 2012.
- [34] K. Tsukada and M. Yasumura. Activebelt: Belt-type wearable tactile display for directional navigation. In N. Davies, E. Mynatt, and I. Siio, editors, *Proc. UbiComp*, volume 3205 of *Lecture Notes in Computer Science*, pages 384–399, 2004.
- [35] Y. Vazquez-Alvarez, I. Oakley, and S. A. Brewster. Auditory display design for exploration in mobile audio-augmented reality. *Personal and Ubiquit. Comput.*, 16(8):987–999, 2012.
- [36] D. Wagner and D. Schmalstieg. First steps towards handheld augmented reality. In *Proc. ISWC*, pages 127–135, 2003.
- [37] S. Wei, G. Ren, and E. O’Neill. Haptic and audio displays for augmented reality tourism applications. In *Proc. HAPTICS*, pages 485–488, 2014.