

Blind People Interacting with Large Touch Surfaces: Strategies for One-handed and Two-handed Exploration

Tiago Guerreiro¹, Kyle Montague², João Guerreiro³, Rafael Nunes¹,
Hugo Nicolau¹, Daniel Gonçalves³

¹LaSIGE, Faculdade de Ciências, Universidade de Lisboa, ²University of Dundee,

³Instituto Superior Técnico, Universidade de Lisboa / INESC-ID

tjvg@di.fc.ul.pt, kmontague@dundee.ac.uk, joao.p.guerreiro@ist.utl.pt, rnunes@di.fc.ul.pt,
hmicolau@fc.ul.pt, daniel.goncalves@inesc-id.pt

ABSTRACT

Interaction with large touch surfaces is still a relatively infant domain, particularly when looking at the accessibility solutions offered to blind users. Their smaller mobile counterparts are shipped with built-in accessibility features, enabling non-visual exploration of linearized screen content. However, it is unknown how well these solutions perform in large interactive surfaces that use more complex spatial content layouts. We report on a user study with 14 blind participants performing common touchscreen interactions using one and two-hand exploration. We investigate the exploration strategies applied by blind users when interacting with a tabletop. We identified six basic strategies that were commonly adopted and should be considered in future designs. We finish with implications for the design of accessible large touch interfaces.

Author Keywords

Blind; Large Surfaces; Tabletop; Strategies; Exploration.

ACM Classification Keywords

H.5.2. User Interfaces: User-centered design, Auditory (non-speech) feedback. Input devices and strategies, Interaction styles.

INTRODUCTION

In recent years, touchscreens have become ubiquitous in society, mostly due to their success in smartphones. Despite being an inherently visual technology, touchscreen devices are growing in popularity among blind users. The built-in accessibility features, such as Apple's VoiceOver or Android's Talkback, allow users to explore and control the device by providing audio feedback for touch actions (*Explore by Touch*). While these accessibility features have been shown to work for small touchscreens, it is not yet understood how practical they are for larger touchscreen interfaces such as interactive tabletop surfaces. The compact form factor of the mobile devices allows blind

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.

ITS '15, November 15-18, 2015, Funchal, Portugal

© 2015 ACM. ISBN 978-1-4503-3899-8/15/11...\$15.00

DOI: <http://dx.doi.org/10.1145/2817721.2817743>

users to acquire spatial references from the physical corners and edges of the device. However, with large interactive tabletops and public displays it is not always possible to obtain these tactile reference points. As a result, exploration tasks become challenging. Furthermore, smartphone interfaces generally present a structured and populated grid of elements that can be used as cues to find the intended target; these are often not available in larger touch surfaces.

Previous work has proposed interface overlays that support non-visual target acquisition on tabletops. These include: linear lists of targets, combined touch-and-speech, and nearest neighbor with voice guidance to target [17]. While these solutions optimize the tasks of locating and counting



Figure 1. One blind participant using our interactive tabletop setup during a trial

targets, it is not clear whether participants were able to gain and leverage a spatial understanding of the interface.

This paper presents a user study with 14 blind participants aimed at understanding *how blind people interact with large touch surfaces* using the *Explore by Touch* approach. Motivated by the two-hand exploration behavior of blind users with physical objects [14] and braille reading [3], alongside the emergence of such approaches in touchscreen

interactions [13], we devised a two-hand alternative of *Explore by Touch*. To support simultaneous bimanual input from two hands, we developed an experimental screen reader using multiple sound sources to provide audio feedback. We combined the detailed touch interaction traces from the tabletop with annotated and coded observations of the recorded sessions, to identify and understand the strategies applied by blind participants with both one and two-hand interaction methods.

The contributions of this paper are: firstly, a user study examining tabletop **interaction strategies** of blind users; secondly, we define **touchscreen interaction metrics** to both identify the presence of, and rationale, for exploration strategies; and thirdly, we propose a set of **implications for the design of accessible large touchscreen interfaces**.

RELATED WORK

The related work reviewed in this section is three-fold: first, we discuss research on novel input and navigation techniques seeking to improve touchscreen accessibility for blind and visually impaired users; second, we address attempts to take advantage of non-visual multimodal feedback to improve spatial understanding; third, we present applications that explore the use of large touch surfaces for visually impaired people (maps, shapes and objects).

Touchscreen Accessibility

Touchscreen devices such as smartphones and tablets have built-in accessibility features (e.g. Apple's VoiceOver or Android's Talkback) that allow visually impaired users to consume the content mainly through auditory feedback. These solutions rely both on gestures and direct exploration, where users drag their finger around the screen while the element in focus is read aloud by the system. Research has tried to improve these interfaces by leveraging gesture-based interaction, including understanding the preference and performance of both new and reference gestures [16], presenting new text-entry methods [12, 21] and browsing the interface using multitouch actions [15]. Other solutions took advantage of multitouch to input text, either using *Braille* (e.g. [2, 19, 24]), supporting two-handed exploration in *QWERTY* keyboards [13] or presenting new keyboard layouts [4].

Multimodal Feedback

While auditory feedback still dominates touchscreen interfaces for visually impaired users, other researchers have explored the use of haptics either alone or together with audio to provide a better spatial understanding. For instance, Giudice et al. [9] provide non-visual access to graphical information (graphs and shapes) on tablet devices by triggering tactile feedback through a single vibrating motor when an onscreen visual element is touched. Goncu et al [10] had a similar goal, but tried to take advantage of the multitouch capabilities of touchscreens, by allowing the simultaneous use of two fingers that could be used to define a line and hear the intersected shapes through sonified 3D

audio. The tactile feedback was provided through vibrating motors attached to the fingers to allow the user to determine the elements positions and their geometric properties. Manshad and Manshad [18] presented a haptic glove that sends vibrations to each finger in order to direct the user to a graph on a grid, in the context of basic algebra.

Nevertheless, these techniques are targeted at very specific domains and do not generalize to all touchscreen interfaces.

Tabletops and Applications

While most research focused on small touchscreens, the knowledge about the accessibility features and research on touchscreens reduces considerably in larger touchscreen interfaces such as interactive tabletop surfaces.

In a first attempt to improve non-visual access to large touchscreens beyond *Explore by touch*, Kane et al. [17] proposed three novel input techniques to enhance the performance of tabletop interactions by blind users. Their proposed solutions decreased the error rates and times taken to complete tabletop interactions by removing the need for users to perform absolute spatial exploration. For example, *Edge Projection* linearized the targets from the fixed 2D positions on the surface into lists of targets that could be explored by sliding a finger along the horizontal or vertical edge of the screen. With this interface, it is possible for users to produce a mental model of targets from their sequence within the lists, as they would for menu navigation on a PC. This approach has strong advantages, as users do not need to accurately maintain a spatial location for each target; they can rely on the reading order of targets to relocate them. In the *Neighborhood Browsing* technique, the whitespace of the screen is divided between the sparse number of targets, and users can request voice guidance to the actual target location from these expanded regions. Voice guidance is also used within the *Touch-and-Speak* [17] condition, where the user can ask the system to read aloud the list of targets, then request directions to its screen position. While both of these interfaces require the user to perform spatial exploration, they are assisted with this interaction and led directly to the desired location.

A practical application that is most frequently associated with large touch surfaces for visually impaired people is interactive maps. Starting with NOMAD [22], several projects relied on tactile map overlays augmented with auditory feedback (detailed review in [7]). Brock et al. [6] later compared a classical raised-line map against an interactive map in a large touch surface with raised-line overlays and auditory output. While efficiency and user satisfaction improved significantly with the interactive map, effectiveness was dependent on the users' individual traits.

In the current study, we observe the natural interaction behaviors applied by blind users when presented with large touch surface exploration tasks. Through this study we uncover the underlying strategies adopted during both one and two-hand exploration. This evidence will support the

development of technologies designed to ameliorate the barriers to non-visual interaction on large touch surfaces.

USER STUDY

This study focuses on identifying and understanding the strategies used by blind users to explore interactive large touch surfaces. We use a laboratory evaluation of tabletop exploration tasks, i.e. *locating*, *relocating* and *counting* targets, and describing the relative position between two targets. We then coded the video and interaction traces to develop an understanding of the characteristics of users' strategies and exploration patterns.

Research Questions

This user study aims to answer, on large touch surfaces, how do blind users:

1. ...acquire targets with no prior screen knowledge?
2. ...leverage prior screen knowledge when acquiring targets?
3. ...spatially relate multiple objects on screen?
4. ...gain an overview of all screen content?

Participants

Fourteen legally blind participants, three females, were recruited from a local social institution. Participants' age ranged from 23 to 62 (M=44.5, SD=12.1). Three participants were left-handed and five participants were experienced with touch phones. None reported having severe motor or hearing impairments. Three participants were congenitally blind. One participant lost sight in the last year. All others were considered legally blind for at least five years. All were screen reader users.

User Interfaces

We collected user interactions with the tabletop surface using our experimental screen reader (based on the current solutions available for mobile touchscreen devices i.e. *Talkback*¹ and *VoiceOver*²). The system was able to capture off-screen hand tracking [20], allowing us to identify which hand corresponds to the touch *id*, recorded by the tabletop. All users interacted with both one and two-hand conditions:

One-Hand is the *de facto* interaction technique for non-visual exploration of touchscreen devices as used by *VoiceOver* and *Talkback*, on mobile devices. Participants could interact by dragging their finger on the screen and the system would read aloud the name of the object as they entered it. The system would also perform a "click" sound when a participant's finger exited an object. Targets could be selected by performing a double tap gesture anywhere on the surface; the last target interacted with was then selected.

Two-Hand extends the basic functionality of the one-hand condition to provide support for simultaneous bimanual interactions. Independent speech feedback was provided for each hand. Based on the previous research investigating

multiple sound sources [11] and non-visual exploration of interactive maps [5, 8], we restricted the number of simultaneous interactions to just one finger per hand. The optimal two-voice setup of the *Text-to-Speeches* spatial audio framework was used to support multiple sound sources [11], each hand was mapped to a specific voice (male or female) and location (left or right ear) to aid interaction feedback distinction.

These interface conditions allow the participants to explore the tabletop interfaces in a similar manner to tactile exploration. Moreover, by minimizing the restrictions on their interaction methods, we are able to observe their spontaneous behaviors and expressions of strategies [25].

Apparatus

This study was conducted using a Frustrated Total Internal Reflection (FTIR) tabletop computer with a 58 x 76cm screen dimension, a Kinect stereo camera positioned above the table, and stereo over-ear headphones were used to provide audio feedback (Figure 1). We used the TACTIC framework [20] to enable multitouch input on the tabletop surface and support hand/finger tracking. We developed a stimulus application to generate tabletop interfaces within a 9 x 12 grid (as used in [17]). We used sparse target layouts to produce screen locations with little or no possible interaction feedback, to both give the illusion of a freeform layout and observe the participants' behaviors under these challenging conditions. Participants were not informed that the targets were distributed in a grid layout or the number of targets in it. Target locations were randomly generated for each new task; all targets were 6cm x 6cm, to fill the cell. The stimulus application captured start and end timestamps, and the screen layout including *target names* and *locations* for each study trial. The interaction data was combined with the screen layout and saved to a log file.

Procedure

Users were invited to the laboratory to participate in the study. Before undergoing the tasks, participants were introduced to the interactive tabletop, provided with a short demonstration, and given time to try each interface condition for 10 minutes. Participants stood facing the tabletop, wearing headphones, as shown in Figure.

The following experimental tasks were used to expose interaction strategies and capture performance measurements for each of the two interface conditions:

1. *Locate*. Given the name (e.g. *Shoe*) of a target, participants were asked to explore the interface to find the target and select it. A single layout with five targets was generated, and participants were asked to locate three targets (one at a time). The second target was only asked once the first one was selected; and the same for the third one.
2. *Relocate*. Having already complete the *locate* task for three targets, participants were then asked to relocate the same three targets in a random order.

¹ <http://developer.android.com/design/patterns/accessibility.html>

² <https://developer.apple.com/accessibility/ios/>

3. *Count*. Participants were given a target name (and a new layout), and asked to count the number of instances of that target within the interface. The number of instances was restricted to a range of 3-5, and the overall number of objects onscreen was of 10. Unlike Kane *et al's* count task [17], each of our target instances had the same name and required the user to maintain awareness of the visited targets.
4. *Relate*. Participants were given the name of two targets and asked to state which target was topmost or leftmost. This required the participants to locate each of the named targets before they could answer the question. As with the *count* task, participants performed the *relate* tasks with different layouts with a screen density of 10 objects (this included the two named targets). To ensure that the *relate* tasks presented a consistent challenge to all participants we defined the following constraints to positioning of the two stimulus targets: in the vertical *relate* task, the targets are at most separated by two rows and at a distance of four columns or more; similar constraints applied to the horizontal *relate* task.

Participants selected the targets by performing a double tap gesture on the intended target. The system would start the task timer from the moment the participant touches the screen, and stop the timer when the task is complete and the participant removes their hands from the screen. The *one-hand* and *two-hand* interface conditions were counterbalanced across our participants. Once the participant performed all of the tasks for an interface condition, they were asked to complete a Single Ease Questionnaire [23] to obtain their perceived ease of use, along with their overall system opinions.

Design and Analysis

The study was a within-subjects design with one independent variable, interface condition. Each participant performed three *locate*, three *relocate*, two *count*, and two *relate* tasks for both interface conditions. We analyse user performance based on the following metrics:

Time Taken: Participants were instructed to remove their hand(s) from the tabletop when they completed each task; the time was then calculated from the difference between the participants' first and last interactions within the task.

Error Rate: An error occurred within *locate* or *relocate* tasks when the participant was unable to find the requested target. If the participant gave an incorrect answer for the *count* task, or wrong position for the *relate* tasks.

Coverage: Using the 9 x 12 grid locations defined above, we count the number of visits participants make to each location, per trial. We also compute the *New* and *Revisited Coverage* for between task analyses of *locate* and *relocate* trials. These values are used to quantify the levels of prior screen knowledge for the subsequent trials.

In addition to the aforementioned performance metrics, we analyzed the strategies that emerged from the exploration tasks. Logged data allowed us to generate animated gif files of the all screen interactions made by each participant (Figure). Effectively, we could replay the users' interactions for each task. These animated traces were then used to characterize and code exploration strategies applied for each task. Video recordings were used to support the coding where users' actions were unclear from the interaction traces. The strategies analysis comprised the coding of touchscreen interaction traces along the seven features shown in Table 1. To verify the reliability of the code set, we refined the characteristics through several phases, similar to the approach applied by Anthony et al. [1]. Initially, two researchers independently coded two sets of 10 randomly selected videos and traces, followed by discussion of the coding dimensions used. In the next phase, the researchers independently coded a set of 40 randomly selected videos and traces (14% of the dataset), the inter-rater reliability was then computed across the seven dimensions and the Cohen's kappa agreement was $k=.89$ ($SD=.08$). The remaining traces were then coded using the same schema independently by the two researchers. The full dataset of interaction traces, used for the coding, are publicly available at <https://goo.gl/FIUqCU>.

<i>Name</i>	<i>Description</i>
<i>Started from the edge</i>	Interaction traces that start from an edge of the screen; this also included the ones that started within the outer cells of the screen.
<i>Structured Movements</i>	Interactions where the trace movements appeared to have a form of organisation and arrangement.
<i>Reference Points</i>	Using other objects as relative location markers to help guide the exploration. One example is shifting direction towards a target when receiving feedback on another object.
<i>Screen Symmetry</i>	Traces where there is a horizontal or vertical axis of symmetry.
<i>Hand Symmetry</i>	Participant's hand movements were proportional in direction and distance.
<i>Simultaneous Input</i>	When the participant uses both hands simultaneously to explore the screen.
<i>Equal Roles</i>	If both hands perform the same operations.

Table 1. Features used to code the interaction traces, and definitions for each.

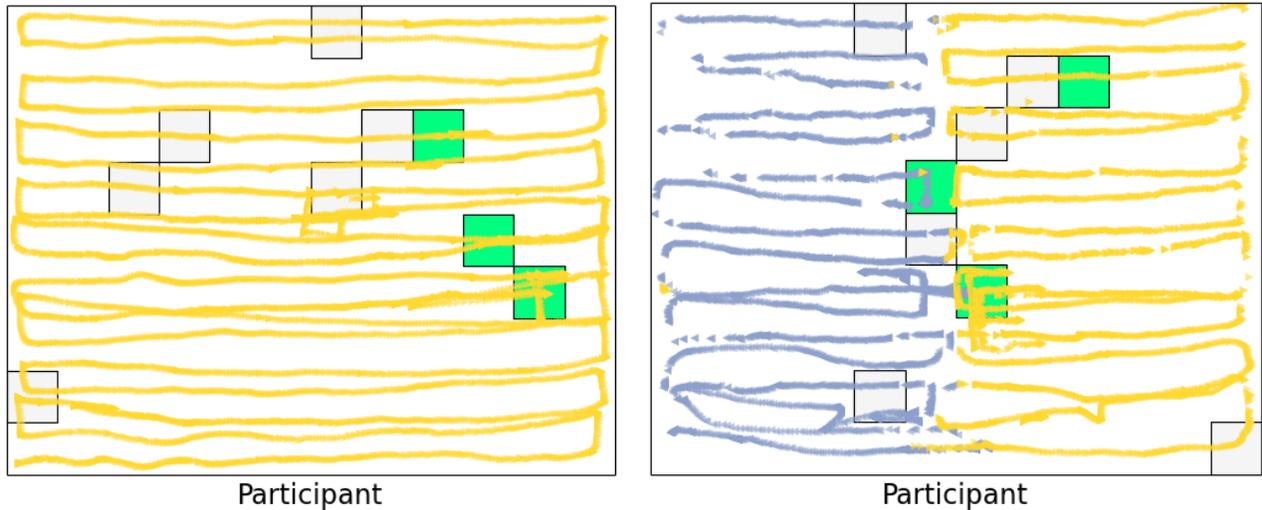


Figure 2. Interaction traces during one of the tasks showing a Path Scan strategy. Left: the participant performs a Count task with the One-Hand condition, Right: the same user performs the same task, but this time with Two-Hand. “Participant” indicates the location of the participant in relation to the tabletop.

FINDINGS

Analysis of the coded characteristics revealed the following usage patterns:

- In 85.71% of *count* tasks, independently from the condition, participants begin their explorations from the edges of the screen, and 55.36% used structured movements, ensuring the best possible coverage and exploration of the space.
- In 66.37% of all *Two-Hand* tasks, the participants used equal roles for each hand. Equal roles were most common in *count* tasks, with their use 89.29% of the time and lowest in the *relocate* tasks with only 38.10%. However, *relocate* tasks have a *simultaneous hand use* of 28.57%, making it less likely that both hands had equal roles in those tasks.

The coded interaction traces were clustered based on their similarity using the features outlined in Table 1. We will now discuss the characteristics and behaviors of the observed strategies, and how they were applied to the tasks.

Strategies

Originally, four strategy clusters were defined from the grouping of the features and manual inspection of the traces: *Path Scan*, *Focused*, *To-The-Point* and *Freeform*. The researchers then discussed the variance within these clusters and opted to divide the *Freeform* to include the *Freeform Symmetry* and *Finger Trailing* variations as their own strategies, giving six exploration strategies.

Path Scan exhibits similarities to the structured scanning paths already applied by screen readers and switch access assistive technologies. Users would start from the faraway edge or corner of the screen, and then drag their finger horizontally until that row has been covered. At which point they would drop a row and return along the horizontal axis in the opposite direction. This action is repeated until the

user either locates the desired target, or reaches the end of the screen. When using *Two-Hand*, users leveraged the second hand by dividing the screen and shared the exploration between the hands. Users would mirror their moves, resulting in a well-structured, symmetrical trace, moving through the screen space row by row. Overall, participants applied this strategy 23.81% of the tasks for *One-Hand* and 30.06% with *Two-Hand*. Figure 2 illustrates a typical *Path Scan* trace in both the *One-Hand* and *Two-Hand* interface conditions.

Focused represents an extensive unstructured search of a small sub-section of the screen where the user believed the target to be. We observed this strategy in situations where the participant had prior knowledge of the target they were looking for. Specifically, the technique occurred during *relocate* tasks and, in some cases, *locate* tasks. However, this could only occur in *locate* tasks when the user had knowledge of that target from a previous *locate* task. Using their dominant hand only, the participant would start the touch exploration not from the edge of the screen, but in the

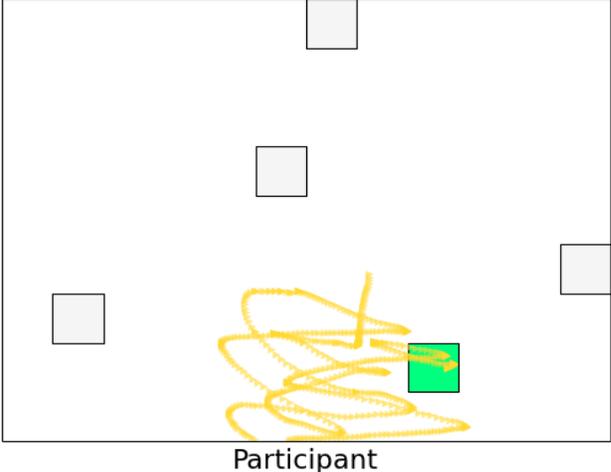


Figure 3. Focused strategy interaction traces of a Relocate task.

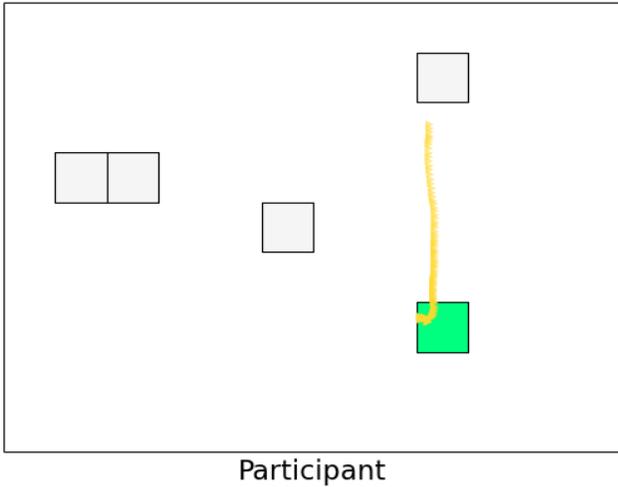


Figure 4. Interaction trace of a Relocate task presenting a To-the-point strategy.

area where they remembered hearing the target before. The movements traces were very close together, and would often overlap several times in a small area. If the participant was unsuccessful in locating the target within the initial sub-section, they would increase the search area slightly and continue with this technique - confident that the target was close by. While participants could explore with two hands simultaneously in the *Two-Hand* condition, they only performed this type of exploration with a single hand. Overall, participants applied this strategy 9.52% of the tasks for *One-Hand* and 5.95% with *Two-Hand*. Figure 3 illustrates a trace applying a *Focused* strategy.

To-The-Point. As with the *Focused* search, this technique relied on participants having prior knowledge of target locations. Again, using the dominant hand, participants would start in an area of the screen close to the target. However, unlike the *Focused* technique, participants had no need to trace and retrace their steps in close proximity of the target. They appeared to know the exact location and could make very structured moves directly to it in a short amount of time. Again, within the *Two-Hand* condition participants did not use the second hand for this strategy. Overall, participants applied this strategy 24.4% of the tasks for *One-Hand* and 14.29% with *Two-Hand*. Figure 4 illustrates the *To-The-Point* strategy, again only a single hand is used, and thus, the traces are the same for the *One-Hand* and *Two-Hand* interface conditions.

Freeform. We regarded this approach as more of an anti-strategy, where participants' movements appeared to be erratic, unpredictable, and without any discernable pattern. Within these traces, we see participants moving around the screen with no real understanding of what has been previously explored. When an object is heard, it is quickly revisited and then the unstructured search continues. Interestingly, in some cases participants were able to retain the spatial locations from this technique, and successful

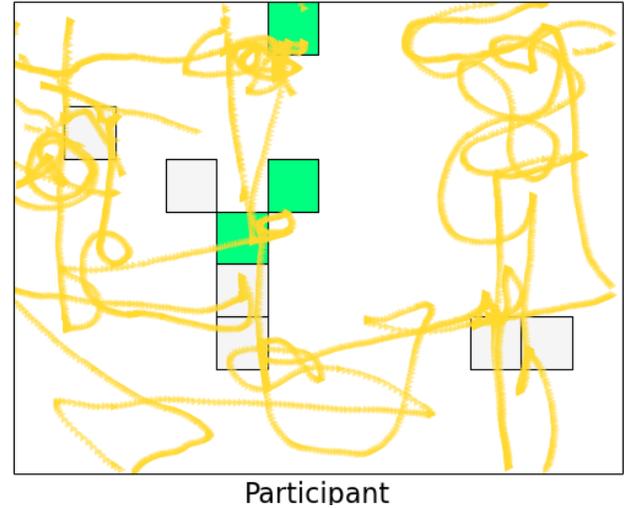


Figure 5. Interaction traces of a Count task showing a Freeform strategy.

apply them in the subsequent *locate* and *relocate* tasks using the *To-The-Point* strategy.

Overall, participants applied this strategy 29.17% of the tasks for *One-Hand*. One example of *Freeform* is show in Figure 5. Only interactions from the *One-Hand* condition were coded as *Freeform*, within the *Two-Hand* we refined this classification to include two variants: *Freeform Symmetry* and *Trailing Finger*.

Freeform Symmetry. As the name implies, this *Two-Hand* variant of the *Freeform* strategy (Figure 6), leverages simultaneous input and hand symmetry to explore the screen in an unstructured manner. While some examples of this technique could be confused for the two hand *Path Scan* approach, they still exhibited an unpredictable nature of movement through the screen. Combining synchronized zigzags down the screen with loops and twists, resulting in horizontal symmetry of the screen. Participants applied this strategy in 15.77% of the tasks.

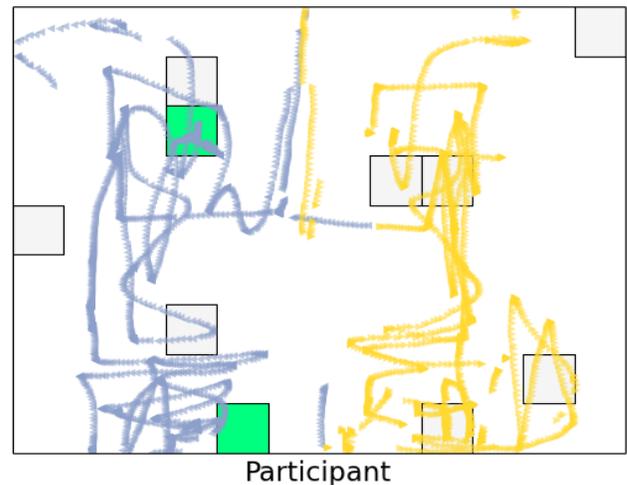


Figure 6. Interaction traces during a Relate task showing Freeform Symmetry strategy.

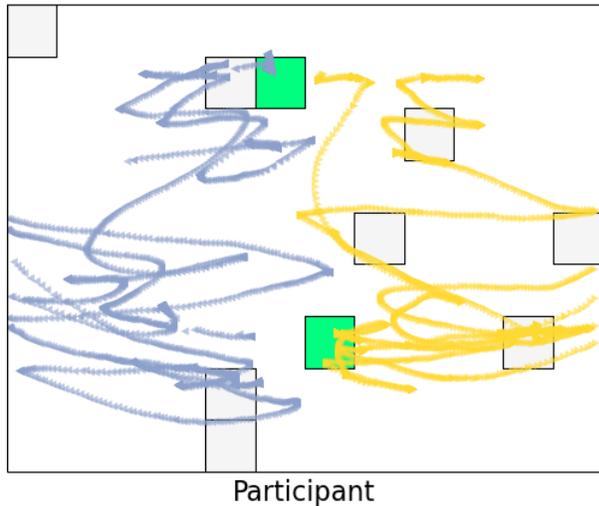


Figure 7. Interaction traces of a Relate task showing a Trailing Finger strategy.

Trailing Finger. With this variation of the *Freeform* strategy, participants used the second finger to search areas in parallel. Both hands shared similar movements but with a horizontal offset between them. Each hand would take in turns to be the lead as the participants moved through the screen. Again, traces were unstructured and fingers moved quickly between different areas of the tabletop. Participants applied this strategy in 29.17% of the tasks. *Figure 7* shows the exploration trace where the user applied *Trailing Finger*, as captured by the *Two-Hand*. However, to understand this strategy the aforementioned animated gifs are needed, to see how the trace was created.

Quantitative Results

In what follows, we present a quantitative analysis for each task and describe how they fit with the strategies employed.

Finding targets without previous knowledge is inefficient. In the first task, participants were asked to search and select a target. This was done with three targets, in sequence, using the same layout and onscreen target disposition. All participants were able to identify the location of the target and only one user failed to accurately select it in the *Two-Hand* condition. They took an average total time to complete the task (search and select 3 targets) of 130.4 (SD=121.8) and 121.7 (SD=61.7) seconds in the *One-Hand* and *Two-Hand* conditions (n.s., Wilcoxon, $Z=-.345$, $p=.730$), respectively. Considering the first target only, when they had no previous knowledge of the layout, participants took an average of 62.1 (SD=82.4) and of 47.84 (SD=36.9) seconds (n.s., Wilcoxon, $Z=-.220$, $p=.826$), respectively. These results suggest that non-visually searching for a target on a large surface, without previous knowledge of its location, is a time consuming task.

The most common strategies applied by the users when performing searching for targets using *One-Hand*, were Path Scan (14.29%), Focused (19.05%) and To-The-Point (30.95%). 33.33% of the searching trials in this condition

revealed a *Freeform* strategy. When two hands were supported - participants applied more structured searches (*Path Scan*, 26.19%) and decreased their use of focused searches (no *Focused* occurrences; *Straight-to-the-point*, 14.29%). 35.71% of the trials were dominated by an unstructured *Trailing Finger* strategy and 21.43% revealed no particular strategy (*Freeform*).

The high level of focused searches in the *One-Hand* setting derives from the design of this task: users performed the three trials sequentially in a grid with the same layout enabling them to acquire knowledge of target disposition for later recall. They were able to leverage that knowledge most of the times they had previously encountered it. Conversely, when using two hands, participants did not build on that knowledge. This could be either because they felt comfortable with the strategy they were using or because they retained less information due to the increased amount of feedback that was provided in this setting.

Considering just the first trial, the number of participants with a structured strategy was the same in both conditions (4). Other participants also opted for dividing the screen and applying a Symmetric strategy (unstructured); this happened only with *Two-Hand* (4 participants). Both structure and symmetry enabled participants to perform a wider coverage of the screen in their explorations with two hands as shown by the total percentage of screen areas covered (*One-Hand*, $M=70.0\%$, $SD=22.5\%$; *Two-Hand*, $M=89\%$, $SD=12.0\%$; Wilcoxon, $Z=-2.971$, $p=.003$). However, due to a less focused strategy with *Two-Hand*, a significant difference was found in how people leverage previous coverage, i.e., *Unique Area Coverage*, ($Z=-2.103$, $p=.035$), between *One-Hand* ($M=42.5\%$, $SD=18.1\%$) and *Two-Hand* ($M=31.2\%$, $SD=12.5\%$). By using two hands, participants seem to have a better coverage of the screen but also to revisit more areas. Notice that using *Two-Hand*, participants cover a larger area and take approximately the same amount of time as when using *One-Hand*.

People can leverage previous screen knowledge. Users were provided with reusable screen knowledge in the locate tasks (the tasks were done in sequence over the same layout) and in the relocate ones, where they had to find previously identified targets. In the first trial, the users cover, in average, approximately, half of the screen in both conditions (*One-Hand*, $M=51\%$, $SD=31\%$; *Two-Hand*, $M=53\%$, $SD=38\%$; Wilcoxon, $Z=-.596$, $p=.551$). In the *One-Hand* condition, as mentioned above, the participants seem to get a good knowledge of the screen, applying focused searches when they have previously visited the targets. The amount of new screen area covered is significantly different from the first ($M=51\%$, $SD=31\%$) to the second ($M=11\%$, $SD=12\%$) tasks (Wilcoxon, $Z=-2.512$, $p=.012$) which then stabilizes at a low level of new coverage. With *Two-Hand*, this effect is only visible from the second (New coverage, $M=30\%$, $SD=30\%$) to the third ($M=6\%$, $SD=14\%$) task (Wilcoxon, $Z=-2.358$, $p=.018$).

Overall, the participants were able to gain knowledge of the targets on screen through prior explorations, then leverage that knowledge of the target locations by applying the *focused* strategy in the final task, in both conditions.

When people were asked to relocate targets they had located, they revealed to have a good notion of the whereabouts of the target, mostly applying a *To-the-point* (*One-Hand*, 66.7%; *Two-Hand*, 42.9%) or a *Focused search* (*One-Hand*, 19.1%; *Two-Hand*, 23.8%). In cases where participants lost the spatial awareness of the target position, they resorted mostly to a *Path Scan* strategy (*One-Hand*, 9.5%; *Two-Hand*, 11.9%) and only few trials showed no particular strategy (*One-Hand*, 4.8%; *Two-Hand*, 2.4%).

These results are corroborated by the percentage of total area covered in the three relocate trials (*One-Hand*, $M=43\%$, $SD=30\%$; *Two-Hand*, $M=7\%$, $SD=14\%$) that is significantly smaller than the one in the three *locate* tasks (*One-Hand*, $M=70\%$, $SD=22\%$; *Two-Hand*, $M=89\%$, $SD=12\%$), in *One-Hand* (Wilcoxon, $Z=-2.534$, $p=.019$) and even more so with *Two-Hand* (Wilcoxon, $Z=-3.297$, $p=.001$). The ability to leverage previous knowledge then translates in the overall performance to complete the tasks: users are significantly faster at relocating targets than at finding them in the first three searching tasks both in the *One-Hand* ($Z=-1.977$, $p=.048$) and *Two-Hand* conditions ($Z=-2.354$, $p=.019$). No significant differences were found between *One-Hand* ($M=68.2$, $SD=63.4$, seconds) and *Two-Hand* ($M=66.9$, $SD=81.6$, seconds) in the time to complete the relocate tasks (Wilcoxon, $Z=.910$, $p=.363$).

Relating targets requires additional aid. Locating and relating the position of two on screen targets showed to be a demanding task. Participants did not correctly relate the targets 46.4% ($SD=30.78\%$) and 28.6% ($SD=37.8\%$) of the times in the *One-Hand* and *Two-Hand* settings, respectively (a minor significant effect was found, Wilcoxon, $Z=-1.890$, $p=.059$). They took an average of 149.6 ($SD=81.4$) and 95.5 ($SD=41.8$) seconds to complete these tasks with *One-Hand* and *Two-Hand*, respectively. Using two hands revealed a significant effect on how fast people could find and relate both targets (Wilcoxon, $Z=-1.977$, $p=.048$). The main reason for this difference is that a second hand on the screen enabled the participants to maintain the position of a first target while searching for the second one. Even in the *One-Hand* condition, some participants used a second hand hovering the screen to *mark* where this target was. The strategies applied before locating a first target was very similar to the ones in the first task (*locate* a target) both in the *One-Hand* (*Path Scan*, 25.0%) and *Two-Hand* (*Path Scan*, 28.6%; *Trailing Finger*, 35.7%, *Freeform Symmetry*, 21.4%). No significant differences were found in the amount of coverage the participants did with both (*One-Hand*, $M=75.8\%$, $SD=18.9\%$; *Two-Hand*, $M=80.0\%$, $SD=17.5\%$; $t(13)=-.790$, $p=.444$).

Lack of understanding of visited space. Participants were asked to count the number of targets with a certain name on

the screen. This task had two main requirements: perform a complete coverage of the screen and be aware of previously visited targets. They took an average time of 126.5 ($SD=54.8$) and 131.6 ($SD=77.1$) seconds (Wilcoxon, $Z=-.031$, $p=.975$) and had an error rate (answered with an incorrect number of targets) of 82.1% ($SD=31.7\%$) and 64% ($SD=36.3\%$) in the *One-Hand* and *Two-Hand* conditions, respectively. No significant differences were found between the two conditions regarding error rate.

Revisiting the aforementioned requirements, participants' performance show a slightly better coverage of the screen (Wilcoxon, $Z=-1.712$, $p=.087$, minor effect) when using two hands ($M=93.6\%$, $SD=6.3\%$) than with one hand ($M=85.2\%$, $SD=17.5\%$). Conversely, participants showed to visit target cells repeatedly and to do it significantly more so when using two hands (*Visits per Target Cell*; *One-Hand*, $M=5.7$, $SD=3.4$; *Two-Hand*, $M=9.6$, $SD=4.7$; $t(13)=-3.181$, $p=.007$). Although using two hands seems to provide the tools for a better coverage of the screen, at the same time users seem to repeatedly visit the same targets, and unknowingly treat those as new targets. These results are supported by the final counts presented: the participants tended to answer with a number below the actual count (average of -0.5 targets) in *One-Hand*, and a higher number in the *Two-Hand* (average of +0.5 targets) condition.

Participants employed a *Path Scan* strategy 46.3% (*One-Hand*) and 53.4% (*Two-Hand*) of the times. Additionally, when using two hands, participants also resorted to *Trailing Finger* (28.6%) and *Freeform Symmetry* (17.9%). The structured strategies enabled the participants to cover most of the screen, however, the participants have no extra cues to understand if they are revisiting a target, and in several cases, an entire line (Figure 2)

Subjective Feedback

We asked our participants to rate each interface condition using 7-point Likert scales for *ease of use* and *speed of use* upon completing the tasks. Wilcoxon signed-rank tests were applied to the Likert scale results. There were no significant differences between the interface conditions for *ease of use*, $z=-.973$, $p=.331$; or *speed of use*, $z=-1.910$, $p=.056$. When asked to state *which interface condition was preferred*, six participants selected *One-Hand*; six selected *Two-Hand*; and two thought that both interface conditions were similar.

The users that preferred *Two-Hand* referred to the larger dimensions of the tabletop, where the two-hand exploration allowed them to cover a larger area of the screen at the same time. One user that made use of the *Path Scan* strategy (Figure 2) responded, "If you estimate the center of the screen and each finger explores half of it, it is way faster. You just have to pay attention". Another user highlighted the role of two hands to provide a better spatial understanding of the screen and its targets, "which is even more beneficial in *Relate* tasks". The main reason for preferring *One-Hand* was a greater confidence in the auditory feedback, as it could only correspond to that

particular finger. In *Two-Hand*, a few users revealed that sometimes they “*needed to move both fingers alternatively to understand which one touched a target*”.

IMPLICATIONS FOR DESIGN

The results obtained enabled us to devise a set of implications for designing interfaces for large touchscreen interactive surfaces accessible to blind users.

Avoid spatial exploration in sparse grids. Nowadays, blind smartphone and tablet users are able to explore the screen with a *One-Hand* approach, which is also known as painless exploration. Our results showed that such exploration, when performed on a large surface, is anything but painless, particularly if the target density is low. Results showed that blind users may lose large amounts of time exploring areas without receiving any type of feedback or cue to direct them towards their intended goal which is time consuming and frustrating. Enabling users to use two hands to explore the screen did not improve their exploration abilities. These results support the need for interface layers that diminish the spatial load, as presented in [17] and allow users to grasp a relative mental model of the contents being.

Establish and maintain spatial models of screen content. In contrast to what was observed in previously unknown layouts, the users demonstrated efficiency in establishing a spatial model of the screen layout and relocating targets. They were able to leverage their spatial understanding to locate targets within previously explored areas. Looking at the preliminary studies performed by Kane et al [17], users were seen to store their items spatially ordered by relevance. They were also able to retrieve them. This ability can be leveraged in layouts that rely on spatial positioning, for example - enabling people to store and rearrange interface items for later recall.

Enable and provide richer feedback. One shortcoming that stood out in the evaluated settings was the inability for users to understand the state of the system, particularly of the explored areas and objects. Resorting to the hyperlink analogy, where previously visited links are presented differently to the sighted user, touch interfaces are falling short in providing feedback to the user about the history of the interaction taking place. While some actions could be detected automatically and feedback changed accordingly, users should also be allowed to enrich the exploration with their own cues and references (e.g., mark areas as visited or other context-related layers). In this paper, with *Two-Hand*, we gave a first step towards enriching the feedback on exploration, mostly by enabling the user to understand which hand touches each object. The design space of sound and/or other feedback channels (e.g., haptics) can be further explored to augment the information the user receives.

More permissive and flexible touch interfaces. As the size increases, so does the need of the user to explore. One of the limitations faced by blind people is that their options are to slide to the next elements or to “painlessly” explore

the screen by finding and receiving feedback on each item. On the other hand, blind users are used to exploring physical surfaces and sheets of paper in search for objects or Braille dots. Some users were seen to leverage the usage of two fingers to split their screen or to assess relationships between objects. Allowing for more natural explorations, similar to the ones applied on physical surfaces, enables users to employ current abilities and thus improve their efficiency. Further research is needed to understand how to provide feedback without overloading the user.

Prioritized elements should be laid along screen edges. Participants were seen to resort to the screen edges as a starting point for their structured searches. This happened particularly when they were aware of the need to be thorough (Count, started from the edge in 85.71% of the trials). These results reinforce that the most relevant interaction contents should be placed along screen edges, particularly in the corners where the users are likely to start their exploration. Given that the most difficult task was locating targets without previous knowledge, the borders can be used as a bootstrap for creating a spatial model of the screen, including interactive contents that allow for accompanied learning of the layout.

Hand roles should be defined based on the task. When allowed to use two hands, users varied in how they utilized this ability. When looking for a good coverage of the screen, users applied equal roles to both hands (exploration) while when relating targets, users used one hand to mark a position and another one to explore. In other cases, like *relocating* targets with a focused search, users seemed to disregard their second hand. Hand roles seem to strongly relate with both the tasks being performed and the strategies applied and should be further explored creating and augmenting new bimanual interfaces for large surfaces.

Active strategy recognition. Current touchscreen interfaces are static and passive to the user’s behavior. The results presented here revealed interaction patterns that speak clearly about the users’ intentions and strategies (e.g., wandering around a target – Focused Search; starting from the edge and performing a structured search – Path Scan). In light of our results, such interfaces should consider adapting to the users’ behaviors and support the interaction taking place e.g., apply Kane’s [17] nearest neighbors, when a long focused search is detected; or provide relationship feedback when a *Relate* behavior is detected, instead of providing feedback of the independent touches.

CONCLUSION AND FUTURE WORK

Exploring the screen of a smartphone has become a common task for a great number of blind people. The onscreen items are linearized in a way that enables their users to quickly grasp an overview of the screen and build a good understanding of the content. Large touchscreen surfaces bring novel challenges to the exploration task. Not only is the exploration surface larger, but the interfaces built for these (e.g., interactive public displays or maps)

often rely on interactions involving spatial positioning or reasoning. We examined how blind individuals interact with large touch surfaces (with a sparse number of targets) to *locate*, *relocate*, and *relate* targets, along with how well they are able to acquire a good spatial model of all screen contents in *count* tasks. Results revealed that the users employ different strategies depending on their goal but that these coping mechanisms are still ineffective. We investigated whether supporting bi-manual exploration would improve performance – which showed to be beneficial in some tasks, particularly to *relate* targets, and to provide better structure in the exploration task. Future work should build on the identified strategies and their shortcomings to develop effective exploration interfaces for blind people and large touch surfaces.

ACKNOWLEDGEMENTS

We thank Fundação Raquel e Martin Sain in Lisbon, and all the participants of the study. This work was supported by RCUK EP/G066019/1 and the Portuguese Foundation for Science and Technology, under grants UID/CEC/00408/2013, UID/CEC/50021/2013, and INCENTIVO/EEI/LA0021/2014.

REFERENCES

1. Anthony, L. et al. 2013. Analyzing user-generated youtube videos to understand touchscreen use by people with motor impairments. Proc. of CHI, ACM (2013), 1223-1232.
2. Azenkot, S., et al. 2012. Input finger detection for nonvisual touch screen text entry in *Perkinput*. Proc. of *GI '12*, ACM (2012), 121-129.
3. Bertelson, P., et al.. (1985). A study of braille reading: 2. Patterns of hand activity in one-handed and two-handed reading. *The Quarterly Journal of Experimental Psychology*, 37(2), 235-256.
4. Bonner, M. et al. 2010 No-Look Notes: Accessible Eyes-Free Multitouch Text-Entry. Proc. Pervasive '10, Springer (2010), 409-427.
5. Brock, A., et al. 2012. Kin'touch: understanding how visually impaired people explore tactile maps. In CHI Extended Abstracts, ACM (2012), 2471-2476.
6. Brock, A., et al. 2015. Interactivity Improves Usability of Geographic Maps for Visually Impaired People. *Human-Computer Interaction*, 30, 156-194.
7. Brock, A., et al. 2013. Map design for visually impaired people: past, present, and future research. *Médiation et Information-Handicap et Communication*, 36, 117-129.
8. Gardner, J., and Bulatov, V. 2006. Scientific Diagrams Made Easy with IVEO. Proc. of ICCHP, (2006) 1243-1250.
9. Giudice, N. et al. 2012. Learning non-visual graphical information using a touch-based vibro-audio interface. Proc. of ASSETS, ACM (2012), 103-110.
10. Goncu C., and Marriot K. 2011. Gravvitas: generic multitouch presentation of accessible graphics. Proc. of INTERACT, (2011), 30-48.
11. Guerreiro, J., Gonçalves, D. 2014. Text-to-Speeches: Evaluating the Perception of Concurrent Speech by Blind People. Proc. of ASSETS, ACM (2014), 169-176.
12. Guerreiro, T. et al. 2008. A. From tapping to touching: Making touch screens accessible to blind users. *IEEE Multimedia*, 15, 4 (2008), 48-50.
13. Guerreiro, J., et al. 2015. TABLETS Get Physical: Non-Visual Text Entry on Tablet Devices. Proc. of CHI, ACM (2015), 39-42.
14. Hatwell, Y., et al.. (2003). Touching for knowing: cognitive psychology of haptic manual perception (Vol. 53). John Benjamins Publishing.
15. Kane, S., et al 2008. Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. In Proc. of ASSETS, ACM (2008). 73-80,
16. Kane, S. K., et al.. 2011. Usable gestures for blind people: understanding preference and performance. Proc. of CHI. ACM (2011), 413-422.
17. Kane, S., et al. 2011. Access overlays: improving non-visual access to large touch screens for blind users. Proc. UIST, ACM (2011), 273-282.
18. Manshad, M., Manshad, A. 2008. Multimodal Vision Glove for Touchscreens. Proc. of ASSETS, ACM (2008), 251-252.
19. Nicolau, H., et al. 2014. B#: chord-based correction for multitouch braille input. Proc. of CHI, ACM (2014), 1705-1708.
20. Nunes, R., et al. 2015. TACTIC: An API for Touch and Tangible Interaction. Proc. of TEI, ACM (2015), 125-132.
21. Oliveira, J., et al. 2011. Blind people and mobile touch-based text-entry: acknowledging the need for different flavors. Proc. of ASSETS, ACM (2011), 179-186.
22. Parkes, D. 1988. "NOMAD": An audio-tactile tool for the acquisition, use and management of spatially distributed information by partially sighted and blind persons. Proc. of International Conference on Maps and Graphics for Visually Disabled People (1988), 24-29.
23. Sauro, J., & Dumas, J. S. 2009. Comparison of three one-question, post-task usability questionnaires. Proc. of CHI, ACM (2009), 1599-1608.
24. Southern, C. et al. 2012. An evaluation of BrailleTouch mobile touchscreen text entry for the visually impaired. Proc. of MobileHCI, ACM (2012), 317-326.
25. Thinus-Blanc, C., & Gaunet, F. (1997). Representation of space in blind persons: Vision as a spatial sense? *Psychological Bulletin*, 121, 1 (1997), 20-42.