

Many Fingers Make Light Work: Non-Visual Capacitive Surface Exploration

Martin Halvey

Department of Computer and Information Sciences
University of Strathclyde, Glasgow, G1 1XQ, UK
martin.halvey@strath.ac.uk

Andrew Crossan

ITT Group, School of Engineering and Built Environment
Glasgow Caledonian University, Glasgow, G4 0BA, UK
andrew.crossan@gcu.ac.uk

ABSTRACT

In this paper we investigate how we can change interactions with mobile devices so we can better support subtle low effort intermittent interaction. In particular we conducted an evaluation with varying interaction techniques which looked at non-visual touch based exploration of information on a capacitive surface. The results of this evaluation indicate that there is very little difference in terms of selection accuracy between the interaction techniques that we implemented and a slight but significant time reduction when using multiple fingers to search, over one finger. Users found locating information and relating information to physical landmarks easier than relating virtual locations to each other. In addition it was found that search strategy and interaction varied between tasks and also at different points in the task.

Categories and Subject Descriptors

H.5.2 User interfaces: Input devices and strategies (e.g. mouse, touchscreen)

Keywords

Non-visual, touch, multi-touch, exploration

1. INTRODUCTION

Recent innovations in mobile interaction and input technologies have led to huge advances in the usability of mobile devices. Technologies such as capacitive touchscreens have opened the door to engaging and aesthetic interfaces that have contributed to an explosion in their usage and in the range of functionality on offer. In spite of these advances however, the way that people interact with these devices has changed very little, with users still selecting from arrays of onscreen buttons as with a desktop interface. Touchscreens allow the user to interact through a small but high resolution screen and directly manipulate widgets and icons. Despite the ease with which it is possible to scroll and resize objects with quick flicks of the finger, they are poorly designed for many common tasks and restrict users to at most a few small single point cursors on screen. The simple act of clicking a button can be a frustrating process as the device must interpret the user's intention by translating an onscreen finger position (that may obscure the target) to a single pixel position. Phones resort to techniques, such as predictive text models for

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.

ICMI '14, November 12 - 16 2014, Istanbul, Turkey

Copyright 2014 ACM 978-1-4503-2885-2/14/11...\$15.00

<http://dx.doi.org/10.1145/2663204.2663253>

typing tasks, for example, to attempt to work around this issue. Problems with lighting conditions combined with the lack of tactile feedback can affect target selection, and if the user is on the move, focusing visual attention on the small screen can interfere with the more important task of safely avoiding obstacles. These issues in some ways make today's mobile devices more difficult to use while on the move than previous generations of mobile phones that used physical buttons.

More recent developments look to improve this mobile interaction mechanism through the use of new technologies. Wearable devices such as Google's Glass and Samsung's Watch¹ are seen by many as the future of mobile interaction. Interfaces such as Apple's Siri software are exploring speech recognition to improve mobile input. However, despite the wealth of research in the area [1, 3, 11, 12] developing optimum mobile inputs for different scenarios is still an open problem. For example Google Glass² attempts to achieve this through augmenting vision. However, augmenting vision may have problems as it requires special potentially expensive and visible hardware; it may be distracting and may not facilitate low levels of engagement. Glass also relies heavily on voice input, which is potentially error prone and not always appropriate. An alternative to visual interaction would be to shift to more continuous multimodal forms of interaction that are more suited to use on the move, while pushing the boundaries of currently available touch technology to its full potential to allow more expressive interaction.

In this paper we will begin to look at one method of fundamentally altering the 'on the move' interaction mechanisms to better support subtle, low effort, intermittent interaction. We begin to examine the process of shifting to non-visual feedback which facilitates variable levels of engagement. We promote closed loop interaction by investigating different non-visual touchscreen input techniques for exploring an audio space. In particular we compare single-finger point, multi-finger point, and multi-finger area cursor interactions for basic search and spatial awareness tasks. The rest of the paper is organised as follows. The next section outlines related work. This is followed by an outline of the experiment conducted, including apparatus, tasks, interaction techniques and procedure. This is followed by an outline of the results that have a mix of qualitative and quantitative results. These results inform a discussion session, which is followed up with conclusions and future work.

2. RELATED WORK

With the rapid advances in the technologies being integrated into mobile devices, the way we interact with these devices is starting to change to better support 'on the move' interaction. In this

¹ <http://www.bbc.co.uk/news/technology-23961692>

² www.google.com/glass

section we will outline the evolution of some of these technologies, with particular attention being paid to touch and non-visual interaction. With the advent of capacitive touchscreens, popularised initially by the iPhone, gesture has now become a dominant modality for input. Capacitive touchscreens allow basic, general purpose functionality like target selection and scrolling to be performed through direct manipulation quickly and effortlessly. Before the iPhone, resistive touchscreens with a stylus mediating the contact were the norm for on the move interaction. However, the whole ethos of capacitive touchscreens was that they were designed to allow users to interact directly using their fingertips. The phone gathers data from a range of capacitive sensors distributed over the screen and reduces these data to one or more single pixel points of contact. This suits the traditional desktop style interactions where target selection from a range of visible targets is the common task. However, we believe that by reducing the data from the capacitive sensor array to single point cursors, much useful and expressive input potential is lost. Further to this, fingertip interaction on a touchscreen introduces its own problems. Siek et al. [19] describes the ‘fat finger’ problem where targeting on a touchscreen can be difficult as the user’s finger obscures the target as they touch it on the screen. There are also issues caused by taking the users contact area on the touchscreen and translating it to a single cursor point on the screen. The system must take a distributed and potentially moving contact area and infer the user’s intent. Rogers et al. [16] note that the centroid of the contact area is often used to determine screen position when targeting on a touchscreen as it is relatively simple to calculate, and demonstrates how to improve this through knowledge of the approach of the finger. The targeting issue is often illustrated through typing tasks on an onscreen keyboard. Typing on a touchscreen has been shown to be slower and more error prone than physical keyboards with the lack of tactile feedback often cited as the main contributing factor. These targeting issues become more problematic on smaller screens such as mobile phones, with many keyboards resorting to predictive text models to help manage the high error rate. Despite these limitations touch technologies are a growing area of interest as they are currently expanding beyond the traditional phone interactions. For example Nintendo have recently released their Wii U³ console that uses a controller with inbuilt capacitive touchscreen. From a mobile perspective, there has also been much interest in touchscreens for non-visual interactions in other areas, for in-car interfaces for example [4].

2.1 Extending the use of touch input

Much of the research looking to improve touchscreen interaction has looked to either improve the accuracy of targeting or improve the management of errors. Baudisch [2] examines different cursor control and display techniques for managing the fat finger problem, particularly for devices with very small screens. One interesting and novel theme of Baudisch’s work is his exploration of back of the device interaction. Here they entirely avoid obscuring the screen by interacting with a separate surface. Many of the interaction techniques explored, however, remain very similar to traditional situations despite the different requirements with the visual channel required for use. Williamson and Murray-Smith developed the Stane interface which allows the user to tap and scrape different textures on the device to interact [14]. The different textures also allow the device to be explored entirely

non-visually with scraping gestures being classified by a neural network. Lyons et al. [13] developed Facet, a multi-display wrist worn system consisting of multiple independent touch-sensitive segments joined into a bracelet. Facet allows users to control how applications use segments alone and in coordination. Applications can expand to use more segments, collapse to encompass fewer, and be swapped with other segments. Yang et al. created Magic Finger [23], which is a small device worn on the fingertip, which supports always-available input. Magic Finger senses touch through an optical mouse sensor, enabling any surface to act as a touch screen. This inverts the relationship between finger and surface, as the finger is the instrument rather than the surface. Kane et al. [10] present three new access overlays with the intention of improving the accessibility of large touch screen interfaces, specifically looking at interactive table tops. Two of the proposed techniques were faster than Apple Voiceover⁴ and were preferred by users. Applications of these overlays include board games and maps. Touch interaction with table top surfaces in particular still utilise visual feedback and can do so in interesting ways. Much of the work presented thus far looks at providing new devices or in the case of Baudisch using a new part of the device, in the work presented here we concentrate on utilising a wider range of the hand for gesture interaction using already existing devices. In a similar vein Wobbrock et al. [22] found for table tops that in many cases visual feedback was required. They conducted an evaluation to try and ascertain user preferences for gestures on table top surfaces. They found that users rarely care about the number of fingers they employ, that one hand is preferred to two, that desktop interactions strongly influence users’ mental models, and that some commands elicit little gestural agreement. There are many examples of novel touch interactions for table tops that use visual feedback. Rock & Rails [20] combines three shape gestures with traditional touch based gestures to increase control, avoid occlusions and separate constraints in 2D manipulation tasks. The bubble cursor [9] is a target acquisition technique based on area cursors where the cursor dynamically resizes its activation area depending on the proximity of surrounding targets, so that only one target is selectable at any time.

2.2 Exploiting Capacitive Sensors More Fully

As outlined in the introduction and mentioned in the previous section the potential for capacitive surface interaction has not been fully explored, and as such they present many opportunities for new interactions. However, there has been some movement towards exploring this area. Sato et al. [17] show how by augmenting everyday objects with capacitive surfaces, specific gestures and the way these gestures are performed can be recognised reliably. One other advantage with capacitive sensing is that it provides a mechanism to sense proximity as well as contact. SNOUT [24] is an interface overlay designed for occasional no-hand or one-handed use of handheld capacitive touch devices. Nose taps, speech-to-text and the accelerometer are used for interaction with the device. Rogers et al. [16] exploit this with finger touch where the interaction includes not only the contact position of the screen but also the direction and angle of the user’s finger as they touch the screen. Multi touch gestures have been pushed heavily as a selling point for Apple’s mobile products, but the fact that visual feedback plays such an important role interaction generally restricts these interactions to the finger

³ <http://www.nintendo.com/wiiu>

⁴ <http://www.apple.com/accessibility/osx/voiceover/>

tips so as not to block the screen. This leads to the rich potentially expressive information about the rest of the user's hand being thrown away as the interaction gets reduced to a few single pixels on the screen.

2.3 “One the Move” Interaction

Beyond capacitive surfaces there is also a host of technologies that now look to solve the on the move interaction problem through other means. One recent area of research explores in-air interactions in mobile situations. Agrawal [1] describes a system that allows users to draw gestures in the air to interact with a mobile device. Kim et al. [11] and Cheng et al. [3] use a wrist based device to detect whole hand gestures allowing users to access a range of functionality through different hand postures. Leap Motion⁵ is a commercially available device that offers in air whole hand gestures in a desktop setting. Much of the research about making information more accessible on the move has concentrated on representing the information through the auditory or tactile channels rather than through vision [5], we look at this in the next section.

2.4 Non-Visual Interaction

From the non-visual perspective, an early example of interaction design for use on the move would be the Nomadic Radio system that allows users to browse a range of content on the move [18]. Previous work has also looked at developing mechanisms for hands-free whole body interaction [6-8]. Results demonstrate how wrist and head rotation along with foot tapping can be sensed and used as input techniques in a mobile setting. Using these techniques, users could search and select information in a mobile setting entirely hands and eyes free combining a few relatively low bandwidth input channel. However, a number of issues were raised about the social acceptability of performing unusual interactions in public [15]. Results show that interactions with a visible device are preferable to gestures where devices are hidden [15, 21]. Costanza [5] avoids this issue using EMG for discreet almost invisible gesturing that allows low bandwidth communication. There are a number of techniques, including haptics etc. that are not outlined here due to space constraints.

In this paper we look to build on the successful approaches outlined in this section by demonstrating far more expressive and continuous control to improve browsing, organising and selecting data. Rather than simple target selection, exploiting whole hand interaction, which we begin to move towards in this work, will allow subtle probing and filtering of the information space to narrow down the information presented.

3. EVALUATION

In this paper, we investigate different input techniques for allowing a user to interact non-visually with a mobile touch surface. A single point of contact is most often used when making touchscreen selections. However, when interacting with a screen, the user can obtain an overview of the interaction space quickly and effectively through a visual glance. If we are working with a touch surface, this glance is not available. A single point exploration of the space can turn searching for the appropriate target into a long and frustrating temporal task. We therefore explore techniques to better support exploration of the space. We first allow multiple points of contact, allowing the user to use more of their whole hand. By distributing the search task among

several fingers, it is hoped that targets will be found more quickly and with less effort. Secondly, interactions with a screen necessarily rely on the finger tips. This allows the user to specify the closest approximation of a point on the screen while minimising the finger obscuration of the target. When interacting with a surface instead of a screen, there is no need to worry about obscuring the target. We therefore allow the user to vary the size of their cursor on the screen by using a smaller or larger contact area. It is envisaged that this will allow a user to approximate the ‘glance’ through using a large contact area with the surface and receiving a wide but unfocused view of the workspace. Small contact areas can then be used for fine grained selection allowing a small but focused area of the workspace to be presented. The goal of this study is to examine how these techniques are used in searching an audio space.

3.1 Experimental Tasks

Three different tasks were used for evaluation. Firstly, locating a target on the surface is a fundamental interaction. Secondly, we test if users can build up a spatial awareness of the targets on the screen by asking them to locate the two closest targets. Finally we ask the users to compare target location with a physical landmark. Details of each of these tasks are outlined below. These tasks are designed to determine whether the user can accurately build up a mental model of the relative positions of the targets in the interaction space non-visually, and how their methods change when allowed to use multiple points of contact as opposed to a single point. Each task involved 3 targets, 3 were chosen as it was the smallest number we could use for some of the tasks which involved comparisons as outlined below.

3.1.1 Locate

Participants are presented with an audio space with 3 targets. The targets are an audio loop of a voice saying ‘Alpha’, ‘Bravo’ and ‘Charlie’ respectively. The task set is for the user to find the ‘Alpha’ target, with ‘Bravo’ and ‘Charlie’ acting as distractors. Searching and locating a target within the space is a basic task that must be performed for a non-visual interface and we would envisage this to be an important task for a large number of interactions with the touch surface.

3.1.2 Closest Targets

The three targets (again ‘Alpha’, ‘Bravo’ and ‘Charlie’) are placed in the audio space. The task is for the user to find the two closest targets. The user must therefore locate the targets and then make judgments about their relative positions. This task tests the user's ability to make spatial judgments between virtual objects. This would be crucial to allow a user to build up a mental model of the distribution of targets within the audio space.

3.1.3 Closest Edge

Three targets are placed in the environment. The task set is to find the target closest to a named edge. An equal number of ‘Top’, ‘Bottom’, ‘Left’ and ‘Right’ judgments are given. Physical landmarks have been shown to provide a useful mechanism for guiding a non-visual search. This task examines a user's ability to make spatial judgments about virtual and physical objects.

3.2 Targets

In each of these tasks, users will be searching for audio targets. Each target has a sound associated with it. Each sound used is a looped recording of a person saying a word from the phonetic alphabet. There are three targets so we use ‘Alpha’, ‘Bravo’ and

⁵ www.leapmotion.com

‘Charlie’ for each of the targets. When the user is not touching the screen or is far away from the targets, no sound will be heard. As a user approaches a target and they get within hearing range (which varies depending on the cursor width), the target sound will start playing. When far away, the audio volume will be low and as the user approaches the target, the audio volume increases reaching a maximum when over the target. The full range of audio volumes allowed by android is used, with the volume at maximum directly over the target and continuously dropping off to zero as the user moves further away from the target. The volume reaches zero at one cursor width distance from the target. Audio latency was not experimentally measured; however, there was no noticeable latency present when interacting.

3.3 Interaction Techniques

The three different interaction techniques implemented, are outlined below and are shown in Figure 1.

3.3.1 Single Point Touch

Here, the user will explore with a single point of contact. This method acts as a baseline control and is similar to non-visual accessible mobile systems like Apple’s VoiceOver⁶. As the user gets closer to a target, the audio volume increases with the sound played at full volume when inside the target. The targets and interaction points are both modelled as circles of fixed size. Values for target radius and interaction point radius were set through pilot testing and at 30 pixels and 100 pixels respectively. Outside a range of 130 pixels, the target cannot be heard.

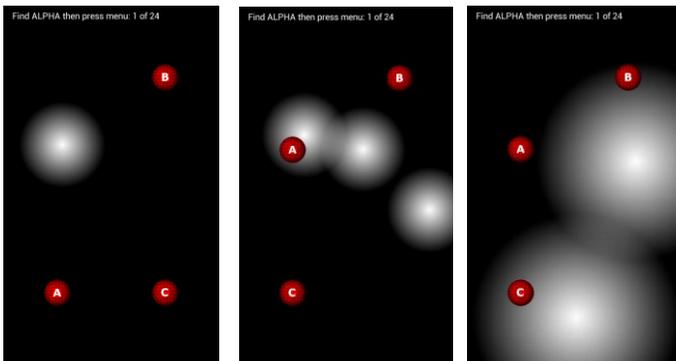


Figure 1: Visual example of the three interaction techniques. From left to right, single touch, multi-touch with 3 points of contact and multi-touch area with 2 points of contact.

3.3.2 Multi-Point Touch

In the second condition, the participant can interact with multiple fingers. Each finger corresponds to an interaction point on the screen. The target size and cursor width are identical to the single point condition. The audio volume of a target is set using the closest interaction point only.

3.3.3 Multi-Area Touch

The capacitive surface allows not only multiple points of contact, but also can be used to gain insight into the contact area of the user’s finger with the screen. Most of this information is hidden in the firmware layer of the phone, however, using standard Android API calls, it is possible to approximate by getting the size of contact area from a touch event using a call to `getSize()`, with the

touch points modelled as circles and the size varying as the radius of the circle. The size value returned is processed further using:

$$\text{Radius on screen} = (40 \times \text{getSize})^2$$

These values were set through pilot testing to allow a wide range of different usable contact areas. This additional information allows the user to control the level of focus in their search with a large contact area allowing a broad area of the audio space to be played to the user, and a smaller contact point being used to present a smaller more focused area of the space to the user.

3.4 Experimental Procedure

The experiment was a 3 x 3 design (interaction x task). The order of interaction technique was counterbalanced to avoid order effects. For each interaction technique block each participant completed the three tasks in the order Locate, Closest Targets, Closest Edge. Participants were given 8 training tasks for each technique/task combination, for the first 4 of those tasks visual and audio feedback was provided and for the final 4 only audio feedback was provided, just like in the experimental tasks. For each interaction technique we measured task completion time, accuracy and a screen trace of user interactions. There were 24 trials for each task giving 216 per participant trials in total. In the ‘Locate’ and ‘Closest Targets’ conditions, the targets were spaced randomly around a 3x5 regular grid on the screen. The same 24 position sets were used for each technique to maintain the same level of difficulty in each condition, but presented in a random order. For ‘Closest Edge’, the target distances from the chosen edge were varied between 80 and 180 pixels with the difference in target distance for each trial ranging between 20 and 60 pixels. Again, the same 24 position sets were used for each trial presented in a random order to maintain the same level of difficulty between trials. Participants completed a NASA TLX workload estimation form after each condition. Participants were instructed to hold the device in their non-dominant hand in portrait mode and interact with the phone in their dominant hand. They could support their arms anyway they wished. Audio feedback was given through headphones. Evaluations took place in a quiet office environment.

3.5 Apparatus

The experimental software was developed in Android. The experimental software ran on a Samsung Galaxy S2.

3.6 Participants

12 participants (10 male, 2 female) aged between 24 and 52 (mean 31.4) took part in the evaluation. All 12 participants had touch screen phones, 9 either owned or had used a tablet, 5 had used a table top and 11 had used touch screen kiosks. All participants were instructed to hold the device in portrait in their non-dominant hand, and to interact with the fingers of the dominant hand. All were right handed and received a £10 Amazon voucher for their participation in the evaluation.

4. RESULTS

As much of the data analysed showed significant differences for Levene’s Test, non-parametric statistical tests were used. The independent variables were analysed using a Friedman’s analysis of variance by ranks, with pairwise comparisons made using a Wilcoxon test. Task and interaction technique were the independent variables.

⁶ <http://www.apple.com/accessibility/osx/voiceover/>

4.1 Errors

Table 1: Accuracy (%) for closest and closest edge tasks for each interaction technique

	Single Point	Multi-touch	Multi-area
Closet Targets	84.85	85.61	84.85
Closest Edge	75.38	75.00	71.21

For the closest targets and closest edge tasks, we calculated the accuracy of the user responses, the results are shown in Table 1. It was found that task accuracy was not effected by interaction technique ($\chi^2(2)=0.743$, $p=0.690$). However, there was a significant difference for task ($z=-5.647$, $p<0.001$), with participants being less accurate at identifying the closest edge.

Table 2: Mean (std. dev.) error x, y and total distance error in pixels for locate task for each interaction technique

	Single Point	Multi-touch	Multi-area
X error	31.923 (33.195)	32.900 (31.340)	32.030 (25.644)
Y error	37.640 (61.064)	33.080 (35.300)	38.540 (37.578)
Error	54.710 (65.359)	52.890 (40.075)	55.740 (38.366)

For the locate task, we measured how close in pixels the user located Alpha to be, we analysed the distance from Alpha in the x direction, y direction and total error. Neither X-error ($\chi^2(2)=0.582$, $p=0.747$) or Y-error ($\chi^2(2)=4.211$, $p=0.122$) was significant. Total error was found to be significant ($\chi^2(2)=7.045$, $p=0.030$), however pairwise comparisons of techniques revealed no significant differences.

4.2 Time to Detection

A summary of the average completion times is in Table 3. Interaction technique was found to significantly affect completion time ($\chi^2(2)=19.670$, $p<0.001$). Pairwise comparisons showed that multi-touch was significantly faster than single point ($z=-3.459$, $p=0.001$) and multi-area ($z=-3.642$, $p<0.001$). Task was also found to have a significant affect ($\chi^2(2)=523.467$, $p<0.001$), with pairwise comparisons showing that all tasks were significantly different at $p<0.001$.

Table 3: Average task time for each task and interaction technique combination, measure is in milliseconds

	Single Point	Multi-touch	Multi-area
Locate	9784.18 (4867.855)	8999.86 (4816.794)	9620.22 (5627.442)
Closest	20881.93 (10023.673)	18584.20 (10852.074)	20036.22 (11419.397)
Closest Edge	11873.79 (6952.276)	11415.40 (6055.607)	12330.604 (7849.816)

4.3 Subjective Workload

Responses to the NASA TLX were evaluated using a repeated measures ANOVA, the entire calculated score rather than the individual scales were used. Interaction technique was not found to be significant ($F_{2,20}=0.590$, $p=0.563$), although the trend was that across almost all differentials that the multi-point touch

technique had the lowest workload. Task was found to be significant ($F_{2,20}=14.546$, $p<0.001$), with locating the two closest points having the highest workload across all differentials. Pairwise comparisons with a Bonferroni adjusted alpha showed that this task was significantly different from locate ($p<0.001$) and closest edge ($p=0.013$). The average responses to each of the scales for task and interaction technique are shown in Figure 2 and Figure 3 respectively.

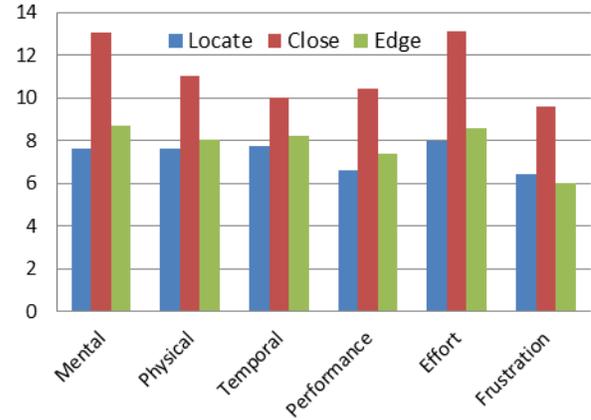


Figure 2: Average TLX responses for each differential per task

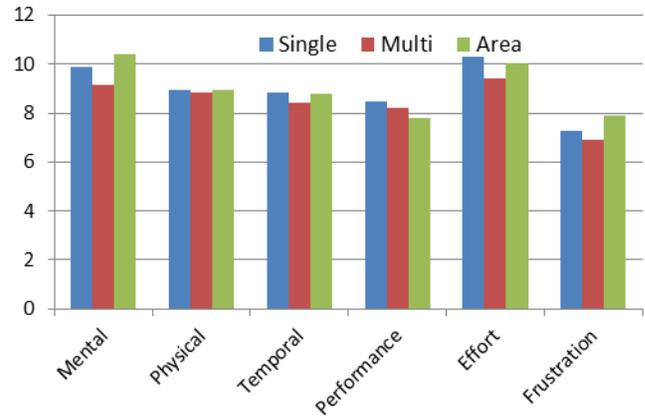


Figure 3: Average TLX responses for each differential per interaction technique.

4.4 Cursor Trace Analysis

Cursor trace analysis was carried out to attempt to identify the techniques that participants adopted using each of the techniques to answer each of the tasks. Here we visualise the cursor traces to categorise the techniques adopted by each of the participants. We also analyse the number of fingers used, the size of interaction points in the area cursor condition and the speed and distribution of the cursors during each exploration.

4.4.1 General Exploration Techniques

The users adopted different techniques for each of the different tasks set. We will discuss these in turn. Locating at least one target was a fundamental task in all of the conditions. In the 'Locate' and 'Closest Targets' conditions, there were no physical cues to guide the exploration. In the large majority of cases, users started at the top of the screen and worked their way down. This was the same regardless of interaction technique. Increasing the number of points available to the user allowed them to adapt their

search. In the single point condition, users generally started at the top-left and zig-zagged right and left downwards until they found a point of interest. This resulted in a search where the users searched first the upper half of the screen and then the lower half in a logical manner (an example is shown in Figure 4 (a)).

When multiple interaction points were available, most users took advantage of the additional interaction points and adapted their search. Two users almost exclusively chose to continue to use one point of contact, even when given the option of using multiple. The general multi-finger technique adopted was to place three or four fingers across the screen and move downwards until a target was found. This had the effect of parallelising the horizontal search between a number of fingers (as shown in Figure 4(b)). Determining the closest targets was generally seen as the most challenging task. This task combined the locate task with a relative position judgment between the different targets, requiring the user to build and maintain a mental model of the target distribution within the workspace. In Figure 5(a) we see how the user zig-zags down the screen in single point mode until all targets are located and then the user moves directly between two targets rapidly to get the answer. In the multi-point mode, the same participant uses the standard multi-finger locate technique then places multiple fingers on the targets to make the decision.

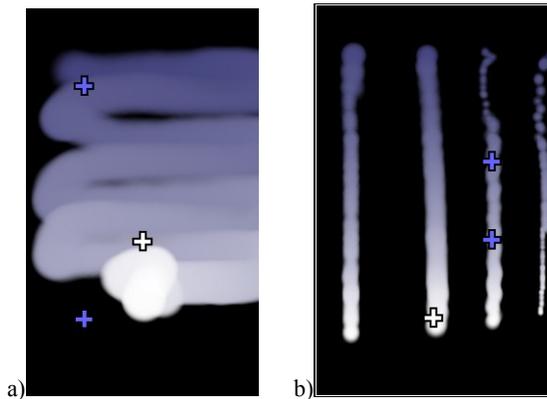


Figure 4: Typical cursor traces in the Locate task using Single (a) and multiple (b) fingers. Darker colours show earlier in the search. The white cross indicates the target to be found

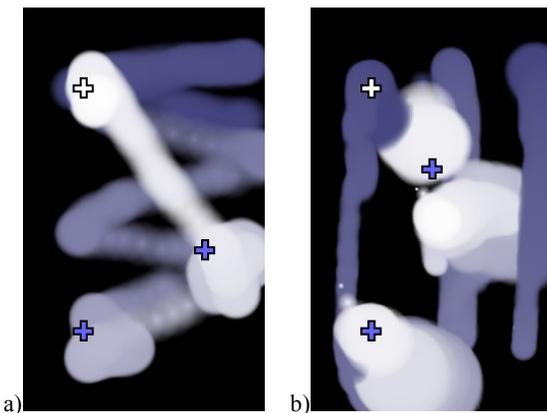


Figure 5: Typical cursor traces in the Closest Points task using Single (a) and Multiple (b) fingers. Darker colours show earlier in the search. The white cross indicates the target to be found

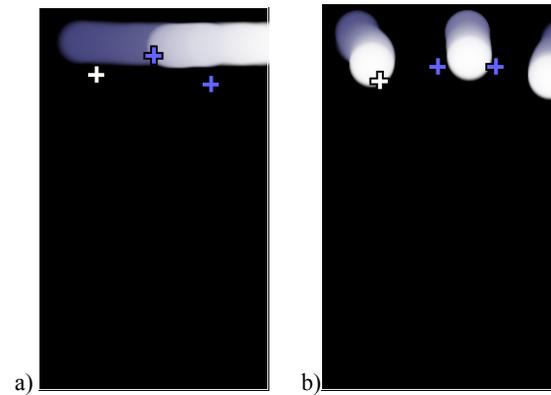


Figure 6: Typical cursor traces in the Closest Edge task using Single (a) and multiple (b) fingers. Darker colours show earlier in the search. The white cross indicates the target to be found

When given a physical cue as to the location to start the search - as was the case in the ‘Closest Edge’ condition - participants used it to guide the search starting point. In the single cursor condition, the technique used almost exclusively was to move back and forth along the appropriate edge rapidly gradually moving further from the edge until points of interest were detected. This was the same technique for horizontal and vertical edges. Again 10 participants used multiple contact points during the ‘Closest Edge’ condition when given the choice. For the top or bottom edge three or four fingers were spread across the screen in a horizontal line at the edge with the user moving away from the edge until they found a point of interest. Orientation affected how multiple fingers were used. For side edge conditions, participants still used multiple fingers but moved them up and down along the edge of the phone as in the single point condition.

4.4.2 Multiple Fingers When Locating

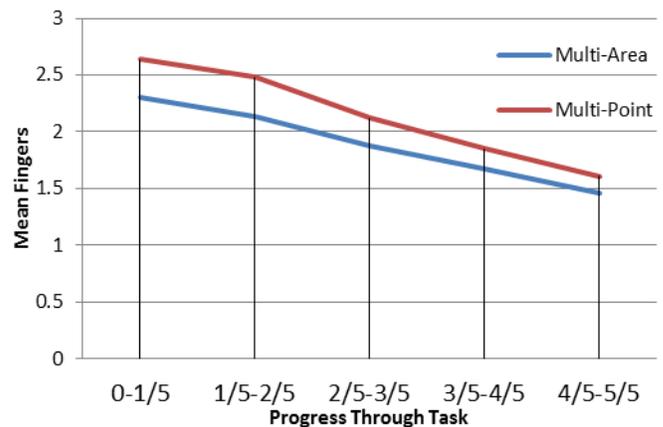


Figure 7: Mean number of fingers on the screen as the Locate task progresses by dividing each task into 5th

If we look at the number of fingers used in the Locate task, we see differing numbers of fingers used as each trial progressed. We split each trial into 5ths, Normalising the time taken to complete the task. 0->1/5th represents the start of the search with 4/5ths -> 5/5th being the end. By looking at the number of fingers used as each task progresses, we see that in the multi-cursor conditions, people used more fingers at the start of the exploration. By using multiple fingers at the start, they were able to search in parallel.

As the target is located, they reduce the number of fingers to perform a finer-grained search over a smaller area (see Figure 7).

4.4.3 Finger Velocity across the Screen

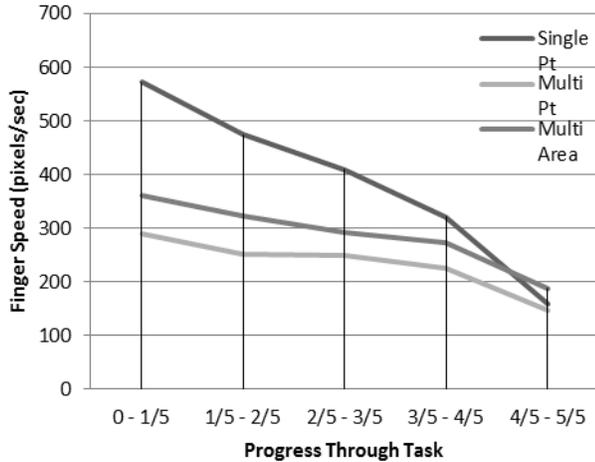


Figure 8: Mean finger speed across the screen for all techniques in the Locate task.

Again if we split the search session into fifths we can see the difference in behaviour of speed of movement across the surface. Looking at mean finger velocity over the screen, we see a large difference between the single-point condition and the multi-point conditions. Figure 8 shows that when multiple fingers were used, the user moved more slowly over the screen. Although the user was able to spread multiple fingers over the screen to search a wider area, they were compensated with the restrictions in the single point mode by moving much faster over the screen.

4.4.4 Cursor Size

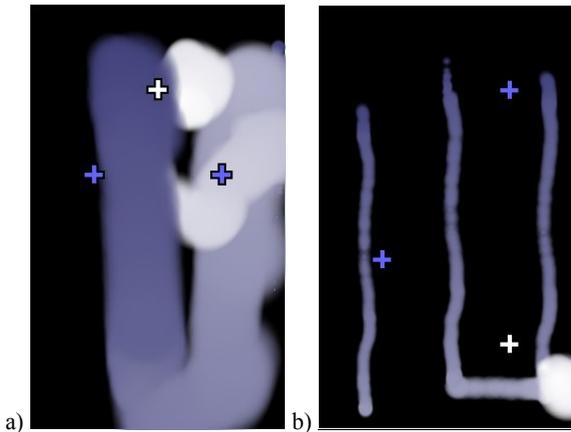


Figure 9: Screen shots of 2 users choosing different area cursor sizes, a) shows a large contact area while b) shows a small contact area.

When examining how participants used the variable cursor width in the ‘Multi-Area’ cursor condition, it was not clear that it was used as part of the search. Participants tended to keep the cursor size approximately the same throughout. Each participant had their own preferred size of cursor; however, this is most likely to factor such as the size of fingers and whether the participants interacted using the very tip or the flat of the finger. Figure 9 shows two extremes, these participants used very

different contact areas with the screen resulting in differing cursor widths, but were consistent between tasks.

5. Discussion

In this experiment we compared three different input techniques for interacting with a capacitive touch surface non-visually. The task set asked participants to locate audio targets on the surface, judge relative distance between targets and judge the position of targets relative to a fixed physical edge. Searching the space and locating the target is the fundamental interaction fundamental to each of these tasks. Using each of these interaction mechanisms, users were able to locate the target to an accuracy of approximately 50 pixels within 10 seconds. Non-visual exploration will always be slower than using visual exploration, however here we target mobile scenarios where vision may be unsuited. Further, this study used an unstructured space with targets appearing randomly throughout the screen. For any system that uses these techniques, we would expect a structured, predictable space that we would expect to reduce time to target. From this perspective, combining the touch surface with audio targets can be seen as a successful mechanism for interaction. We were further able to demonstrate that different techniques were used by participants when single or multi-point interactions were used. For a single finger search, the user zig-zagged left and right down the surface. When using multiple fingers, by far the most common technique was to spread three or four fingers in a line near the top of the screen and move downwards. Results show this method lead to slightly faster search times in the Multi-Point condition than in the Single-Point condition. This could be due to the fact that the search task is being parallelised between the fingers reducing the need to move side to side. This time difference was somewhat offset by the fact that users moved a single point on the surface faster than when the more cumbersome multiple-points were in contact. There was no effect detected on the targeting error.

The area cursor technique was included to allow both a broad view and a focused view of the space. The broad view is an attempt to replace the visual glance that we rely on for everyday interactions with mobile devices. The focused view was intended to allow for more fine-grained interactions. However, from the cursor trace data; there is no evidence that participants used the cursor sizing as an intentional input mechanism. Participant’s maintained relatively constant cursor sizes throughout the study although these varied between participants. In this instance either through a lack of perceived usefulness, or through lack of previous exposure to a novel interaction technique, it was not used by these participants. It is still an open question whether any performance benefits can be gained by this technique. Multiple fingers on the surface served a similar purpose creating a cursor distributed across multiple points on the surface. There is evidence that users placed more fingers on the screen during the initial search phase. Towards the end of the search as the target was located, fewer fingers were used showing transition from a broad search to a narrower focused search

When asked after the experiment, participants stated a preference for the single finger condition. This was a lab-based study where users were sat at a desk with phone in hand. This is a familiar situation for phone usage and the fact that they were able to move their finger over the screen in a posture and manner similar to everyday usage may have influence their opinions. Similar performance was seen in each of the given task for each of the cursor techniques. It remains to be seen whether when mobile, the

ability to ground multiple fingers on the surface will have a steadying effect on the hand and easy mobile interaction. Of the given tasks, judging the spatial relationships between the virtual targets was the most difficult this is borne out by the additional time required to complete the task as well as the NASA TLX results. Participants did however manage to successfully complete the task with a good degree of accuracy (~85% in all conditions). Multiple fingers here reduced the need to move between he targets when located, but did not lead to any performance gains.

6. CONCLUSION

We conducted a study to examine the performance of three different input techniques aimed at eventually allowing non-visual low effort interactions in a mobile setting. Results showed that participants were comfortable searching an audio space using one or multiple fingers on a capacitive surface with multiple fingers showing a significant reduction in time for locating a target. Participants still stated a preference for the more familiar single point of contact interaction. These results along with the analysis of the strategies employed by the participants when using one or multiple fingers on a touch surface will aid designers of mobile or wearable system who consider similar interaction styles. Future work will look to extend this lab study to a more realistic mobile setting with the device located either in pocket or worn on a sleeve or belt. We will also look again at the question of the use of area cursors as this remains an open question. With appropriate training, this technique may still provide a mechanism that will allow both an overview of the space to be gained while still allowing accurate selection. This is the first in a series of studies where we will look to push the boundaries of capacitive touch input. By allowing the user to exploit more of their whole hand, we aim to develop discrete, low effort interactions that will benefit on the move interactions for mobile and wearable devices.

7. REFERENCES

- [1] Agrawal, S., Constandache, I., Gaonkar, S., Roy Choudhury, R., Caves, K. and DeRuyter, F. Using mobile phones to write in air. *ACM MobileHCI* 2011, 15-28.
- [2] Baudisch, P. and Chu, G. Back-of-device interaction allows creating very small touch devices. *ACM CHI* 2009, 1923-1932.
- [3] Cheng, J., Bahle, G. and Lukowicz, P. A simple wristband based on capacitive sensors for recognition of complex hand motions. *IEEE Sensors* 2012, 1-4.
- [4] Colby, R. Considerations for the Development of Non-Visual Interfaces for Driving Applications, Doctoral dissertation, Virginia Polytechnic Institute and State University (2012).
- [5] Costanza, E., Inverso, S. A. and Allen, R. Toward subtle intimate interfaces for mobile devices using an EMG controller. *ACM CHI* 2005, 481-489.
- [6] Crossan, A., Brewster, S. and Ng, A. Foot tapping for mobile interaction. *BritishHCI* 2010, 418-422.
- [7] Crossan, A., McGill, M., Brewster, S. and Murray-Smith, R. Head tilting for interaction in mobile contexts. *ACM MobileHCI* 2009, 6.
- [8] Crossan, A., Williamson, J., Brewster, S. and Murray-Smith, R. Wrist rotation for interaction in mobile contexts. *ACM MobileHCI* 2008, 435-438.
- [9] Grossman, T. and Balakrishnan, R. The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. *ACM CHI* 2005, 281-290.
- [10] Kane, S. K., Morris, M. R., Perkins, A. Z., Wigdor, D., Ladner, R. E. and Wobbrock, J. O. Access overlays: improving non-visual access to large touch screens for blind users. *ACM UIST* 2011, 273-282.
- [11] Kim, D., Hilliges, O., Izadi, S., Butler, A. D., Chen, J., Oikonomidis, I. and Olivier, P. Digits: freehand 3D interactions anywhere using a wrist-worn gloveless sensor. *ACM UIST* 2012, 167-176.
- [12] Lumsden, J. and Brewster, S. A paradigm shift: alternative interaction techniques for use with mobile & wearable devices. *Proc. of the 2003 conference of the Centre for Advanced Studies on Collaborative research*. IBM Press, 2003, 197-210.
- [13] Lyons, K., Nguyen, D., Ashbrook, D. and White, S. Facet: a multi-segment wrist worn system. *ACM UIST* 2012, 123-130.
- [14] Murray-Smith, R., Williamson, J., Hughes, S. and Quaade, T. Stane: synthesized surfaces for tactile input. *ACM CHI* 2008, 1299-1302.
- [15] Rico, J. and Brewster, S. Usable gestures for mobile interfaces: evaluating social acceptability. *ACM CHI* 2010, 887-896.
- [16] Rogers, S., Williamson, J., Stewart, C. and Murray-Smith, R. Fingercloud: uncertainty and autonomy handover incapacitive sensing. *ACM CHI* 2010, 577-580.
- [17] Sato, M., Poupyrev, I. and Harrison, C. Touché: enhancing touch interaction on humans, screens, liquids, and everyday objects. *ACM CHI* 2012, 483-492.
- [18] Sawhney, N. and Schmandt, C. Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM transactions on Computer-Human interaction (TOCHI)*, 7(3) (2000), 353-383.
- [19] Siek, K. A., Rogers, Y. and Connelly, K. H. Fat finger worries: how older and younger users physically interact with PDAs. *Interact* 2005, 267-280.
- [20] Wigdor, D., Benko, H., Pella, J., Lombardo, J. and Williams, S. Rock & rails: extending multi-touch interactions with shape gestures to enable precise spatial manipulations. *ACM CHI* 2011, 1581-1590.
- [21] Williamson, J. R., Crossan, A. and Brewster, S. Multimodal mobile interactions: usability studies in real world settings. *ACM ICMI* 2011, 361-368.
- [22] Wobbrock, J. O., Morris, M. R. and Wilson, A. D. User-defined gestures for surface computing. *ACM CHI* 2009, 1083-1092.
- [23] Yang, X., Grossman, T., Wigdor, D. and Fitzmaurice, G. Magic finger: always-available input through finger instrumentation. *ACM UIST* 2012, 147-156.
- [24] Zarek, A., Wigdor, D. and Singh, K. SNOUT: one-handed use of capacitive touch devices. *ACM AVI* 2012, 140-147.