Making touch-based kiosks accessible to blind users through simple gestures

Frode Eika Sandnes*, Tek Beng Tan, Anders Johansen, Edvin Sulic, Eirik Vesterhus and Eirik Rud Iversen

Faculty of Engineering, Oslo University College, P.O. Box 4, St. Olavs Plass, N-0130 Oslo, Norway

E-mail: frodes@hio.no

Tel: +47 22 45 32 49

Fax: +47 22 45 32 05

Abstract

Touch-based interaction is becoming increasingly popular and is commonly used as the main interaction paradigm for self-service kiosks in public spaces. Touch-based interaction is known to be visually intensive and the current non-haptic touch display technologies are often criticized as excluding blind users. This study set out to demonstrate that touch-based kiosks can be designed to include blind users without compromising the user experience for non-blind users. Most touch-based kiosks are based on absolute positioned virtual buttons which are difficult to locate without any tactile, audible or visual cues. However, simple stroke gestures rely on relative movements and the user does not need to hit a target at a specific location on the display. In this study a touch-based train ticket sales kiosk based on simple stroke gestures was developed and tested on a panel of blind and visually impaired users, a panel of blindfolded non-visually impaired users and a control group of non-visually impaired users. The tests demonstrate that all the participants managed to discover, learn and use the touch-based self-service terminal by completing a ticket purchasing task. A majority of the participants completed the task in less than four minutes on the first attempt.

Keywords: accessibility, touch interaction, audio feedback, self-service kiosk, blind users

1 Introduction

Touch technology has recently become a popular and widely used technology [23]. Users and developers alike have been mesmerized by the appeal of touch interaction. Consumer electronics products, such as the iPhone, have contributed to brining touch to the average user [1].

Touch interaction is a challenge for blind users mainly because current technology does not provide tactile feedback. Efforts have been taken to make both the iPhone and other smart phones accessible. Apple’s VoiceOver and The Eyes-Free Android project are examples where touch is combined with text-to-speech to guide the users. Touch interaction for blind users is currently an active area of research. For instance, Tinwala and MacKenzie showed that stroke based text input, or gestures, inspired by handwriting can be performed without visual feedback on handheld devices [36]. The input can also be enhanced by exploiting multi-touch capabilities such as in the no-look notes system [4]. One challenge with gestures inspired by handwriting is that learning is required. An alternative approach is simple gestures based on directional strokes where each direction controls a choice. Such directional strokes have successfully been applied in Yfantidis and Evreinov's text entry strategy [41] Sánchez and Aguayo's messenger for the blind [24], Sánchez and Maureira subway mobility assistant [25] and O'Neill et al.'s patient information system [19]. A totally different approach could be the use of multi-touch displays for the input of chords [28].

Next, McGookin, Brewster and Jiang showed that MP3-players can be controlled using gestures without visual cues [17], but that there can be problem with short impact-related operations. The same research also showed that an ordinary touch-based MP3-player could be operated successfully with a paper overlay providing tactile cues.

The Slide-rule system [12] employs a general gesture language for portable devices where users can browse lists of options with vertical movements were the fingertip explores linear lists with audio feedback and selections are performed with horizontal flicks of the finger. This strategy relies on absolute positioning of the fingers which is feasible on small devices. The Slide-Rule system also relies on multi-touch gestures which are currently only supported by a few platforms.

The low hardware costs and the fact that the interface is completely designed in software means that touch also has become a popular interaction paradigm for self-service kiosks installed in public locations providing services such as tickets sales [30], Internet access [8], information [33, 34, 37], city guides [11], voting [6], electronic questionnaires [3], photo services [21], banking [20], etc. These machines reduce costly staff and contents can be changed in real time from a central location. However, interfaces based on touch interaction are often visually intensive and ticket vending machines with touch displays have been criticized by Schreder et al. for excluding blind users who are unable to see the display contents [32]. For example, the touch-based train ticket machines deployed in Norway has been criticized for being inaccessible. Although ongoing research is focusing on developing displays that provide haptic feedback [2, 5, 10], few commercial products offer such functionality. To a blind user the touch display is one continuous surface unlike apparatus with physical buttons that provide valuable tactile cues to blind users. The inaccessibility is especially problematic for self-service kiosks that provide essential services to the society such as ticket sales for the public transport sector. In addition, some touch displays are positioned at seat-height so as to be reachable by wheelchair users, but then become very hard to read for tall users with reduced vision who need to lower down into an uncomfortable posture to read the display contents. Displays may also be difficult to read due to insufficient contrast or reflections in the display surface as pointed out by Hagen and Sandnes [9].

Self-service kiosks installed in public locations usually have very simple button based interfaces where users make a sequence of choices. Most users have been trained to understand the virtual button metaphor and understand that the visual representation of buttons affords pushing [18]. A few years ago one could find users who did not know that one should touch the screen. However, the population at large has been educated and once one sees one person touching a screen, others are quick to follow. Currently, most users probably expect any display installed in a public location to be interactive through touch.

Button based interaction requires absolute positioning with the hands. Physical buttons can be felt and spatial motor memory can help users learn the position of buttons without having to resort to visual cues. However, without visual or tactile feedback it is difficult to find an absolutely positioned target. Audio feedback can be used to guide the user, such as the talking fingertip technique [38], where the user searches the display with the fingers to locate absolute positioned targets, but such exploratory techniques are slow if performed in two dimension on a large area. The talking fingertip technique has also inspired several other systems such as Slide-rule where the search is made more effective by narrowing the search from two dimension down to one-dimensional linear lists [12] or circular search, such as employed in the Earspod system [42], which can also be considered a one-dimensional search space. Absolute positioning is less suitable for most kiosk displays as they are comparably much larger than the displays on portable devices where some degree of eyes-free absolute positioning has proven possible, such as in the Slide-rule system [12].

The Apple iPhone combines a mixture of absolute and relative gestures in its voice-over interface. Users explore absolutely positioned items with the fingertips just as with the talking fingertip technique and relative gestures such as one, two and three finger flicks are employed to go up and down lists, taps anywhere on the screen are used for making selections and twisting motions are used for making selections (rotary dial metaphor). Although it has been shown that the iPhone voice over user interface is accessible [22] and used by many blind users throughout the world it is likely that these blind users represent the technologically savvy and curious individuals. The gestural language employed by the iPhone is relatively complex and near impossible to discover without instruction. The text input strategy, for instance, require the user to search the virtual keyboard with one finger and make selections with the other. On the other hand, a public kiosk is quite different to a personal mobile device. First, it is used infrequently, and sometimes it is just used once. Second, the user has less opportunity to explore and learn the interface as one can with a mobile phone. A user interface on a mobile device can be learned at the user's own pace without onlookers in the comfort of one's own home. A public kiosk is often used with bystanders observing one's every move and it is embarrassing to make mistakes. Moreover, current touch technologies mostly support only single touch interaction. Multi-touch capabilities such as that found on the iPhone and the iPad are still quite uncommon.

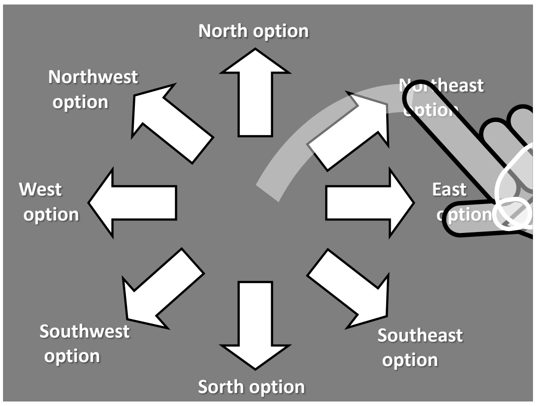


Fig. 1. Visual cues guide non-blind users along absolute paths.

This study is based on the phenomenon that relative stroke motions, often referred to as gestures, do not require visual or tactile feedback as it can be performed open loop. For instance an upwards motion or a downwards motion can be performed anywhere on the display real-estate. Moreover, displays are usually installed with a frame that can be felt with the hands and as such can provide an initial frame of reference. Such frames thus allow blind users to identify the boundaries of the display. Non-blind users, however, are guided by the visual cues presented on the display (see Fig. 1).

One key challenge is that the general population is less accustomed to gestural interaction compared to point and touch. Although gestures have been successfully been used for operating touch-based mobile devices, there are at time of writing no documented study touch-based public kiosks controlled by relative gestures intended for novice infrequent blind users. This study therefore proposes the novel idea of using simple relative gestures for the operation of touch-based public kiosks. Moreover, experimental evaluations are used to demonstrate that if it is possible for general blind users to explore and quickly learn to perform simple gestures and quickly learn to use a simple gestural language.

2 System design

In this study a prototype self-service train ticket kiosk was designed with the goal of achieving what Savidis and Stephanidis call a dual interface, that is, an interface that is accessible to both users with and without vision [31]. The interface had to be instantaneously usable with no need for prior training. In addition, the interface had to support impatient frequent travelers who expect kiosks to provide speedy service.

The kiosk was designed to acquire the following information from the users: desired language, ticket type (single-return), number of passengers, fare type (adult, child, student), the destination and payment type. The path through the system is similar to that of the national self-service train ticket machines in Norway. Note that the path through the system could be greatly optimized. A discussion on self-service terminal path optimization can be found in Thimbleby [35].

Although important, the detailed mechanisms for payment were omitted in this study for two reasons First, payment systems are often standardized in Norway, and separate units on the kiosks. The design of these is governed by regulations where security takes precedence over usability. Second, the authors viewed it as to personal for the participants to use their own payment cards in the trials entering their personal pin number, even on a dummy terminal. Thus we omitted issues of establishing the participants' trust and the bureaucracy of acquiring permission from the Norwegian Social Science Data Services.

The kiosk is controlled through simple directional stroke gestures on the touch display as input, and feedback was provided both visually on the display and via audible speech.

The interface is structured around a set of dialogues, where each dialogue acquires one unique piece of information from the user.

2.1 Gestural input

A gesture language comprising a set of single directional strokes was selected for controlling the input. A rightward stroke signals a *select* while a leftward stroke signals *back* (undo) – both modeless actions. This choice was based on the Western left-to-right reading direction [26]. The remaining six directions, that is, up down and the four diagonal directions were used for making modal selections depending on the particular dialogue.

The user has a chance to explore all the options in each menu before moving to the next menu. It is the most recently selected option that at any time is chosen. The user can also go back to a previous menu by using the back-gesture (left stroke). Note that there is no direct way of going back to the start screen. To go back to the start users need to use the back-gesture several times until they have returned to the start. Alternatively, the users can wait for the session to time-out. Selecting the appropriate time-out duration is a non-trivial task as at one hand one does not want a machine to be locked half way into a session for too long after a machine is abandoned, while on the other hand users with cognitive disabilities need sufficient time to operate the machine. Providing a separate start-over gesture is one possibility. However, the challenge is to design a gesture that is intuitive and easy to remember, while at the same time is sufficiently different from the directional gestures used to operate the machine.

During the design it was noted that some handheld touch devices are implemented using a dragging metaphor where users make a leftwards dragging motion to move rightwards in the information stream and vice versa (see Fig. 2). Here, the direction of the stroke is in the opposite direction of the desired motion. The adopted approach could therefore cause a stimulus response incompatibility for users accustomed to the dragging metaphor.

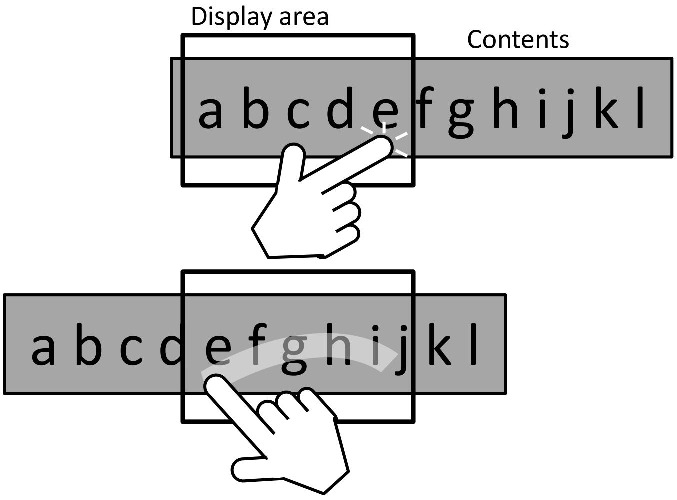


Fig. 2. The dragging metaphor – a leftwards stroke results in a left scroll that moves the user towards the right in the information stream.

2.2 Audio output

The audio feedback was presented using speech. To obtain a neutral and clear set of speech samples a commercial speech synthesis software package (Voxit Budgie Pro) was used to generate the audio. Since the system was tested on Norwegian participants both the textual information and the audio samples were provided in Norwegian. The speech samples comprised three types of audio – *instructions*, *options* and *confirmations*. The instructions were brief and carefully worded. A direct English translation of the instructions are as follows:

> *"This ticket vending machine is operated by dragging the finger in various directions along the display. Six directions are used for making selections, namely diagonally up-left, up, diagonally up-right, diagonally down-right, down and diagonally down-left. When you have made a selection drag the finger to the right to continue. To go back to the previous menu, drag the finger left. To hear all the alternatives you need to drag the finger in the various directions. To start drag the finger right."*

The initial instructions convey details about how to use the kiosk and for each dialogue instructions on the available options in that particular dialogue represented using descriptive phrases. Brief confirmations were provided after each dialogue and a complete summary of the selections made was provided in the last dialogue.

The initial instructions were presented automatically when the user first touches the screen. Similarly, for the final summary screen audio feedback is played automatically. For the remaining screens the audio is activated by the user. Each option is read out as it is being explored. The user explores the various directions until the right option is found. The user can then commit to a selection by issuing the next-gesture. Two distinctive sounds are used to signal that the user is moving to the next screen or the previous screen, respectively.

2.3 Language selection

Fig. 3 shows an example of the dialogues encountered while purchasing a ticket. The first dialogue (Fig. 3.a) presents instructions on how to use the kiosk, namely the purpose of the kiosk and how to operate it using the simple gestures. After reviewing the instructions, or at any time during the playback of the instructions, the user can input a right stroke (select) to move to the next dialogue – the language selection (Fig. 3.b). In this prototype the user has a choice of Norwegian, English, German and French. Note that only support for Norwegian was implemented in the prototype and this dialogue was included to illustrate language selection. By moving the fingers in various directions the user can explore the options. For example, an upwards stroke selects Norwegian. The choice is signaled to the user as the selection is highlighted and played back as “Norwegian”. The selection remains visible while blind users will have to repeat the directional stroke if they want to hear the selection again. Alternatively, the user can input a northeasterly stroke to select English, southeasterly stroke to select German and downwards stroke to select French. Once satisfied the user makes a right stroke to move to the next dialog upon which the selection of the current dialogue is confirmed.

The inclusion of language selection on self-service terminals is generally disputed as some argue for visually presenting multiple languages simultaneously [30], but the inclusion of audio feedback strengthens the arguments for language selection as it is challenging to simultaneously present speech in multiple languages. In a production system careful thoughts need to go into how the language selection mechanism should work. In the current prototype the language selection comes after the basic instructions, while these Norwegian instructions are not accessible to travelers who do not know Norwegian, such as tourists. One practical approach is to display very basic instructions in the main languages on the first screen such as Norwegian, English and Chinese, and then allow the users to select their desired language. Alternatively, the first screen could also be implemented as a continuously running video that visually illustrates how to operate the kiosk with short round-robin audio instructions in multiple languages.

2.4 Ticket details

After language selection the user has to choose the type of ticket, namely single (one-way), return (two-way), seasons ticket or the collection of a pre ordered ticket (Fig. 3.c). The selection is made in the same manner as in the language selection menu, that is, the user moves the finger in one of the six directions to explore the available options. Single tickets are the most common and this is therefore chosen as the default choice. Frequent users can therefore directly select the default by inputting a right stroke. Note that such a dialogue is strongly affected by the ticket and pricing structure for a given train operator in a given country.

In the next dialogue (Fig. 3.d) the user selects the number of tickets. One ticket is the default choice and is represented by an upward motion. Alternatively the user may select 2, 3, 4, 5, or more tickets. The next dialogue (Fig. 3.e) queries the user for the fare type with adult as the default type. Again, various train operators may have totally different fare systems, and we have adopted the one that is used by NSB (Norwegian State Railways), namely adult, child, senior, student, soldier and dog.

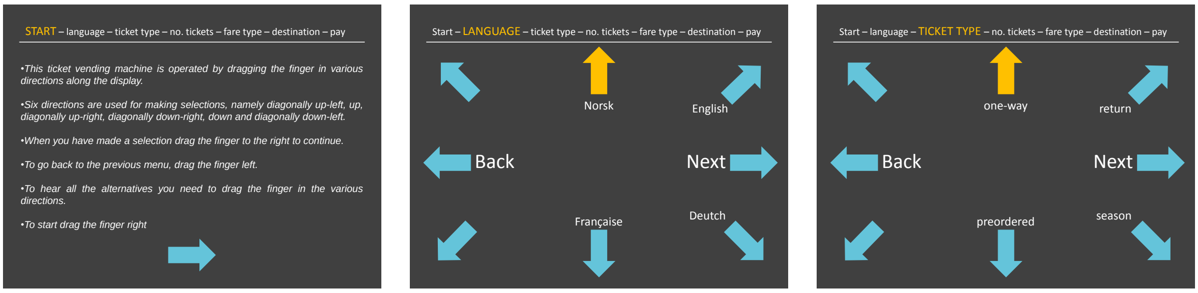
2.5 Specifying destination

The next few dialogues (Fig. 3.f-i) represent what is probably the most challenging phase of ticket purchases, namely selecting the destination. Any non trivial public transportation system will comprise many destinations. For example, the Norwegian rail network comprises several hundred stations. Clearly, it is not possible to present such a large number of destinations in a simple dialogue and the selection has to be done in several steps in some hierarchal approach. Destinations can be organized geographically, alphabetically, according frequency of travelers, etc. Presenting frequently visited destinations, such as Oslo Central Station or Oslo Airport is useful, but was discarded because only six gestures were available for options. Moreover, a geographical organization requires users to have a geographical semantic or spatial understanding of the train network structure, which is an unrealistic assumption. Instead, a selection strategy based on place names was adopted. Note that the place name selection strategy is not completely unproblematic as users need to know how to spell the place name of a destination.

The strategy, inspired by systems such as Yfantidis and Evreinov's adaptive blind text input system [41], Kurterbach and Buxton's marking menus [14], Sánchez and Aguayo's messenger for the blind [24], Sánchez and Maureira's subway mobility assistant [25] and O'Neill et al.'s patient information system [19] works as follows. The user first selects the category of the first letter of the destination where each of the six selection directions represents the categories *abde*, *fghi*, *jklm*, *nopr*, *stuv* and *yøå* (see Fig. 3.f). Note that the letters *c*, *q*, *w*, *x*, *z* and *æ* are not included as there are no destinations beginning with these letters. After making a selection such as *jklm* the user is taken to the next dialogue (see Fig. 3.g) where the user selects between *j*, *k*, *l* and *m* presented in clockwise direction. Next, the user needs to provide a second letter to limit the search – in this instance the user chose between *ma*, *me*, *mj*, *mo*, *my* and *mø* (see Fig. 3.h). In this example, enough information is then provided for the user to select the destination from a limited list, namely, *Mo i Rana*, *Moelv*, *Moi*, *Mosjøen*, *Moss* or *Movatn*. Clearly, the number of selections needed in order to make an unambiguous selection will vary depending on the name as the tree structure representing all the destinations is unbalanced. From an implementation perspective a trie is a suitable data structure for storing destination names [7].

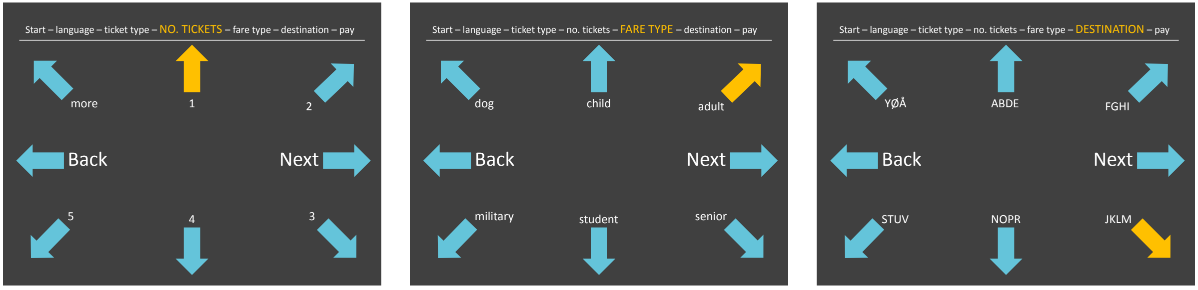
Specifying destinations based on the spelling is effectively a text entry operation. Destination selection is thus often implemented using virtual keyboards. Clearly, virtual keyboards are not suitable for blind users with the current non-haptic technology because of the lack of feedback. Although the proposed strategy is slow it does not require any particular text input skills, knowledge such as keyboard layouts or training [27]. Moreover, it is operated with fewer steps than other techniques commonly used in consumer electronics products, such as the date stamp text input technique studied by MacKenzie [16].

Next, having specified the destination, the user is presented with a confirmation dialogue (see Fig. 3.j) summarizing the selected options, both visually and aurally. Finally, the user is taken to a payment dialogue (Fig. 3.k). The current prototype gives the user a choice between credit card and cash to illustrate the possibility. In practice however, this dialogue is redundant as the terminal will automatically know the means of payment when the user either inserts a credit card into the credit card reader or cash into the notes and coin inlets, respectively.

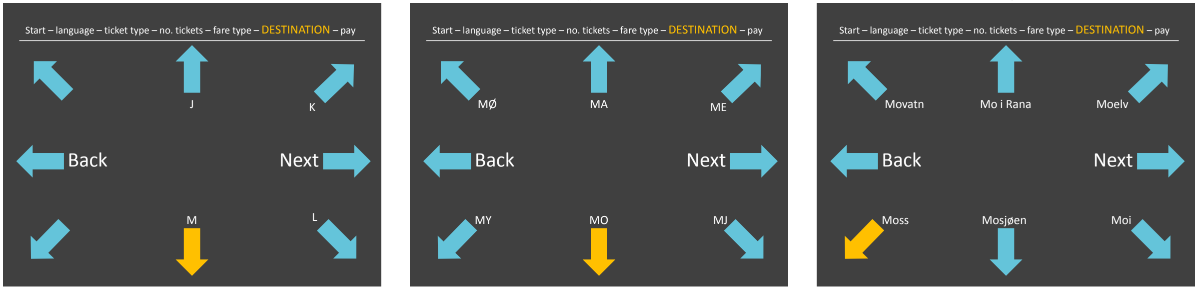


- *This ticket vending machine is operated by dragging the finger in various directions along the display.*
- *Six directions are used for making selections, namely diagonally up-left, up, diagonally up-right, diagonally down-right, down and diagonally down-left.*
- *When you have made a selection drag the finger to the right to continue.*
- *To go back to the previous menu, drag the finger left.*
- *To hear all the alternatives you need to drag the finger in the various directions.*
- *To start drag the finger right*

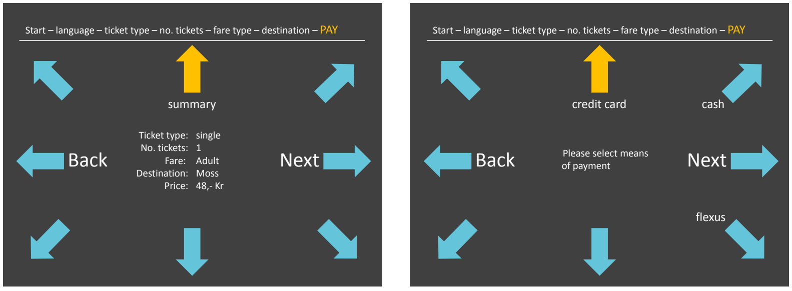
a) Instructions b) Language selection c) Ticket type



d) Number of tickets e) Fare type f) Selecting destination



g) Selecting destination h) Selecting destination i) Selecting destination



Ticket type: single
No. tickets: 1
Fare: Adult
Destination: Moss
Price: 48,- Kr

Please select means of payment

j) Confirmation k) Payment

Fig. 3. Purchasing a train ticket using the self-service terminal (translated).

2.6 Implementation

The prototype was implemented in Adobe Flash and run on a portable personal computer with Microsoft Windows 7 connected to an Irontech capacitive touch sensitive display. This display was mounted into a polystyrene catering case specially cut in an angle to give the sensation of using an authentic self-service terminal when placed on a table (see Fig. 3). Several touch technologies were evaluated including several resistive touch displays, of which few seemed suitable for accurately capturing the gestures. The capacitive display chosen was sufficiently sensitive to effectively capture the single-stroke gestures.

A very simple gesture recognition engine was implemented by capturing the horizontal and vertical display coordinates of the mouse-down event (x_{down}, y_{down}), and the horizontal and vertical coordinates of the mouse-up event (x_{up}, y_{up}). First, a check is performed to see if the gesture is sufficiently long to represent a stroke, namely

$$T^2 < \delta x^2 + \delta y^2 \qquad (1)$$

where T is a threshold representing the minimum distance in pixels that can make up a valid gesture. The threshold T was set to 1/8 of the number of pixels along the diagonal which corresponds to a short distance on the display surface. Moreover,

$$\delta x = x_{up} – x_{down} \qquad (2)$$

$$\delta y = y_{up} – y_{down} \qquad (3)$$

The angle a of the gesture direction is computed as

$$a = \text{atan2}(\delta x, \delta y) \qquad (4)$$

If the angle a is in the range -12.5 to 12.5 degrees it is a rightwards (east) stroke, if it is between 12.5 and 57.5 degrees it is a up-right (northeast) stroke, if it is between -12.5 and 57.5 degrees it is a down-right (southeast) stroke, the range 57.5 to 112.5 degrees represents up (north), while -57.5 to 112.5 degrees represents down (south), etc. For more complex unistroke gestures a slightly more sophisticated strategy is needed such as the 1$ gesture recognizer [40].

Fig. 4. The self-service kiosk prototype implemented using a touch display mounted in a polystyrene case.

The gesture based self-service ticket kiosk was evaluated using a panel of users. These tests are outlined in the next section.

3 Experimental evaluations

3.1 Participants

Three groups of participants were recruited, namely 10 individuals with no visual impairment who served as a control group, 15 individuals with no visual impairment that were blindfolded and a group of 16 individuals with varying degree of visual impairment including blindness. The 25 participants without visual impairment were computer science students at Oslo University College. It can be assumed that these participants are highly computer literate and probably have knowledge about touch interaction. The 16 participants with varying degrees of visual impairment were recruited among the staff and visitors at the Norwegian Association of the blind in Oslo. The level of visual impairment among the 16 participants were in the range of category 2 (severe visual impairment) to category 5 (blindness – no light perception) using the World Health Organization’s definition of visual impairment [39]. Although several of the participants were categorized as blind they had some level of perception to light (category 3 and 4). The level of visual impairment for each individual was not recorded to preserve the anonymity of the participants in accordance with the Norwegian Personal Data Act.

3.2 Task

The participants were asked to purchase one single journey adult ticket to Moss or Sandvika using the self-service kiosk. The participants were instructed to select the Norwegian language option, verify their purchase and pay using credit card. An experienced user would be able to execute this task with just 12 gestures as the

Norwegian language selection and adult single fare are defaults that can be selected with simple leftwards strokes.

3.3 Apparatus

The experimental setup shown in Fig. 4 was used, that is, the touch sensitive display mounted into the polystyrene casing, attached to a laptop computer running the prototype software. The audio was played via the built in speaker on the laptop computer. The laptop computer was located sufficiently close to the touch display for the speech to be clearly audible to the participants.

3.4 Procedure

Each participant was given a brief introduction to the task before being guided in front of the touch display. Of the 25 participants without visual impairments 15 participants were blindfolded. The participants were asked to talk aloud while performing the tests. In addition, total task completion time and the total number of gestures used were recorded.

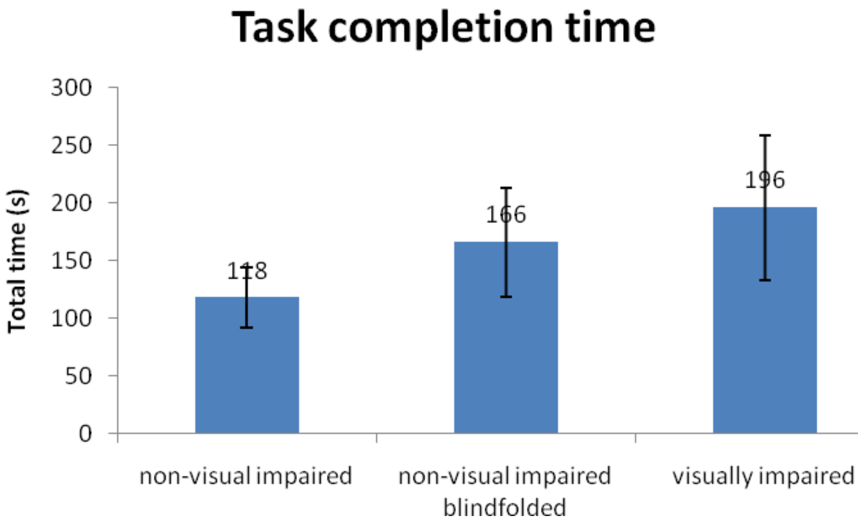


Fig. 5. Median time in seconds to complete the task for the three test groups. Error bars show IQR.

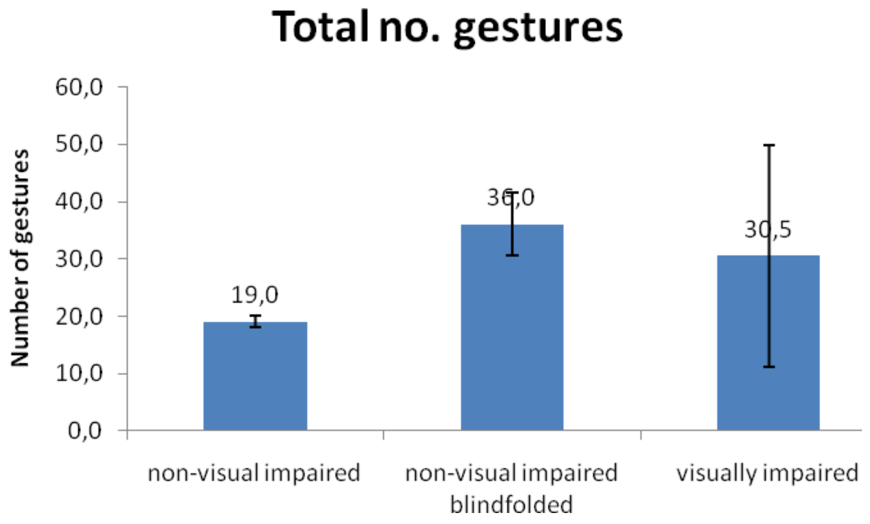


Fig. 6. Median number of swipes needed to complete the task for the three test groups. Error bars show IQR.

3.5 Results

Figs. 5 and 6 summarize the user tests. The graphs show the median time needed to complete the tasks for the three groups, where the error bars represent the inter-quartile range (IQR), and the median number of gestures used to complete the tasks for the two groups, with error bars representing the inter-quartile range. Median and inter-quartile measures for centrality and spread were used as they are robust to outliers, as two of the groups had one participant whose results deviate from the others. The outlier among the blindfolded non-visually impaired and visually impaired groups needed 350 and 690 seconds to complete the task, respectively. Both of the participants who took the longest time were the most impatient. They attempted to rush through the task, but ended up taking longer time. The slowest participant in the control group took 216 seconds to complete the task. This participant took longer than the others because he was very calm and gave a detailed account of his choices during the session.

3.6 Discussion

The experiments demonstrated that all the participants managed to complete the task and interact using the simple gestures without training, including the participant with category 5 blindness. The results show that the differences between the three groups is relatively small, where the participants with no visual impairment used about 2.0 minutes, the blindfolded participants with no visual impartment used about 2.7 minutes and the visually impaired participants used 3.3 minutes to complete the task. Although some of the non-visually impaired participants were blindfolded, they completed the task faster than the blind users. One explanation for this phenomenon is that the non-visually impaired participants were computer science students and were less inhibited and felt free to explore the interface and managed to more quickly adapt to the interface compared to some of the visually impaired users who were primarily older administrative staff with less computer experience. Thus, age and computer experience are factors that are likely to have influenced the difference between the two groups. Note also that several of the participants with no visual impairment reported feeling uncomfortable operating the computer blindfolded. The results show that the control group had the shortest task completion time and the difference between the blindfolded non-visually impaired and those that were not blindfolded demonstrate that the visual cues are important. Moreover, this result demonstrates that the strategy is feasible for both non-visually impaired and visually impaired users.

The visually impaired participants had the largest spread in task completion time among its users. This spread could be caused by the various types of visual impairment. The control group had the smallest spread in task completion time. An one-way ANOVA test reveals that the task completion times for the three groups are statistically different (F(2,38)=3.7;p<.04).

In terms of number of gestures used then the blindfolded non-visually impaired users needed the most gestures, that is, 36 gestures, to complete the task. The larger number of gestures is a result of the users spending more time exploring the interface. Overall, the control group completed the task with the fewest gestures, that is, 19 gestures. As the members of the control group could see the alternatives they did not explore the interface using audio. Next, the visually impaired users needed 31 gestures to complete the task, signaling that they were the more confident than the blindfolded non-visually impaired participants. One reason for this could be that visually impaired users rely more, or fully, on the audio feedback and thus adopts a somewhat different exploration strategy than users that mainly are used to rely on visual cues. Another explanation could be that the visually impaired users may be slower, more cautious, or afraid of errors. Note that the spread is also larger for the visually impaired users than for the users without visual impairments. The spread is very low for the control group suggesting that these effectively did not employ any exploration strategy. This group did not have to explore the various alternatives because they could see them all at once. A one-way ANOVA test reveals that the number of gestures used by the three groups are statistically different (F(2,38)=15.4;p<.001).

The participants who did not rely on visual cues input approximately three times as many gestures as the theoretical minimum (12 gestures). These extra gestures were the result of exploring the various options. The number of gestures needed is likely to decrease if the users are repeatedly exposed to the system. The repetitions will help users memorize the default options and thus make shortcuts.

The control group input just below twice as many gestures as the theoretical minimum (19 gestures). This is because they did not exploit the default options available on some of the screens where they could have gone straight to the subsequent screen. For example, on the screen for selecting the number of ticket the default choice is 1 ticket which is implicitly chosen by going to the subsequent screen. None of the members of the control groups explored the various options, but a few participants made one or two mistakes which is the reason for the variance. Also, only one of the control group participants chose to listen to the summary.

When making their first choice several participants in the control group thought the system had crash as nothing happened once they had selected an option, and it took some time before they realized that they had to explicitly go to the next screen after making a choice. One subject even repeated the same option selection gesture in the hope that this would help him get to the next screen. Another participant attempted to push the words on the navigation bar at the top of the screen. Once the two-step select-next concept was grasped it caused no confusion thereafter. Clearly, the non-visually impaired users, relying on the visual cues, expected to be taken to the next screen immediately after making a choice. Clearly, such an approach will not allow for exploring the options for visually impaired users. Perhaps one way of solving this would be to give explicit visual feedback indicating that they have to select next to continue. This could for instance be implemented as an information bubble popping up close to the finger once an option is selected successfully.

A fraction of time also went into reading the textual instructions on the start screen and it takes about 30 seconds to listen to the audio. There is a lot of information to absorb at once. Perhaps it would be better to present the instructions gradually on successive screen, that is, perhaps first introduce the notion of going forward and backwards, then choosing between one of two alternatives and then finally six alternatives. The user would then learn the procedure as an integral part of purchasing the ticket. Obviously, this may require the sequence of the various screens to be altered. One participant in the control group also suggested using longer arrows as a visual guide for the fingers.

Overall, a majority of the participants managed to complete a ticket purchase transaction in less than four minutes on the first attempt which is an acceptable result. The two outliers needed 5.8 and 11.5 minutes, respectively. Note that the outlier in the visually impaired group had some vision and the long task completion time was a result of not immediately understanding the interface – not because of a lack of visual cues.

When using the interface for the second, third, and fourth time, etc, the total time to conduct a transaction was reduced as the users became more familiar with the interface and know how to interact effectively in order to reach their goal quickly.

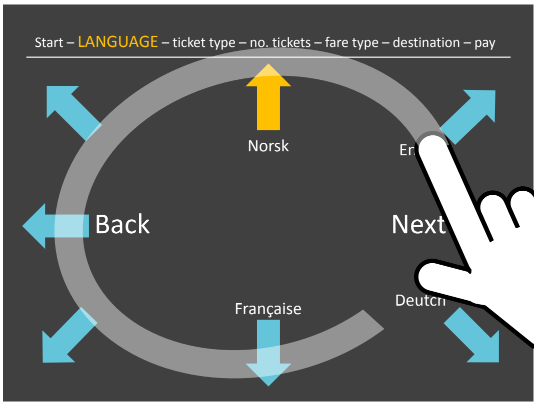


Fig. 7. Some participants attempted to explore the available options with circular gestures.

One interesting observation made during the tests was that several participants among the visually impaired group attempted to explore the options using a circular motion (see Fig 7). Once a mental model of options being available **in** the eight directions was established these users deduced that the options are available on a circle attempted to access the various options with this circular gesture similar to the circular motions used to interact with the wheel-based ipod [29] or Earpod system [42]. This suggests that the proposed approach possibly could be combined with such a pattern. It is computationally feasible to distinguish between circular exploratory gestures and the simple stroke gestures and circular gestures can thus also be implemented and used as relative gestures, that is, they circular exploration can be performed anywhere on the touch-display area.

Generally, the comments made by the participants during test were those of positive surprise with expressions such as “cool” and “really good idea”. Several blind users pointed out that it was the first time they had ever managed to successfully use a touch sensitive display to complete a task. One of the more critical points raised during the tests was that selecting destination is difficult. Opinions varied regarding how the letters should be assigned to directions. Moreover, some participants complained that it was difficult to hear the difference between the spoken letters *M* and *N* during the destination selection phase.

4 Limitations and future work

A key issue, not addressed in this study, is how blind users locate self-service terminals in the public landscape – especially when navigating around unknown territory. One solution that has been proposed is to use some kind of beacon realized with existing infrastructure such as Bluetooth [15]. Other aspects of a kiosk can pose a challenge such as finding the money inlet or credit card slots. This problem is analogous to the challenges of making bar codes accessible to blind users [13].

A different perspective is whether self-service kiosks are needed at all [35]. One alternative strategy would be for the user to plan, prepare and purchase tickets in the comfort of their own homes with their preferred assistive technologies and then use a mobile device to assist them once travelling. Again, unplanned situations occur occasionally in which one has to purchase tickets just before boarding.

Another possibility that could improve the accessibility of self-service kiosks for visually impaired and blind users is for accessibility settings to be associated with the travelers’ credit card or electronic travel card. Once inserted, or swiped, the kiosk could adopt to the user. However, this strategy has privacy challenges that would need to be resolved.

During the tests some of the participants presented the wish that the self-service kiosk also double up as an information kiosk by providing audio information about the track from where the train is departing and the departure time. Often, such information has to be acquired from separate information boards that are often not accessible to blind or visually impaired travelers.

An increasing number of personal devices are employing gestural interaction on touch surfaces, and the cost of these devices is continuously dropping. There are some signals to suggest that the gestural languages are converging on certain alphabets. It may be possible that these devices may change our expectations for, and knowledge of, gestural interaction, and that this also may be exploited in future self-service touch-based kiosks.

The menu items in this prototype were limited to 6 items, with the exception of the text input that relied on hierarchal text input. However, the real world often has complex ticket pricing, such as geographical price zones and off-peak/peak prices. One approach accommodate more than six option is to employ a two-level ticket hierarchy which allows for 36 options, however, the effectiveness of such schemes need to be verified through user testing.

5 Conclusions

This study has explored the use of simple gestures to make touch-based self-service kiosks accessible to blind users, without sacrificing the user experience for users without visual impairments. A simple prototype was implemented based on a simple gesture language allowing the user to select, undo and explore various options. Feedback was provided using synthetic speech. The experimental evaluations showed that most participants managed to complete a ticket purchasing task in less than four minutes on the first attempt – including blind and blindfolded users as well as a control group of non visually impaired users. The results show that some mechanism is need to signal to non-visually impaired users that they explicitly need to go to the next screen after making a choice.

The gestures are easy to learn and simple to implement. The tests showed that several participants attempted to use circular motions to explore the available options. Future work will therefore include combining the directional gestures with circular gestures for accelerated exploration of options. Our results show that gestural input can be used on self-service terminals by users with impaired vision. Such interfaces should rely on simple and relative gestures. Moreover, the gestural language should be simple and consistent.

Acknowledgements

The authors are grateful to Ole Petter Braathen at Activum for providing a license to the Voxit Budgie Pro speech synthesis software used in this project and Jørn Jensen at Irontech who lent us the touch display used during the tests. We are also grateful for the kind help from the Association of the Blind, Norway, for help with recruiting participants for the tests.

References

[1] Arruda-Filho, E. J. M., Cabusas, J. A., and Dholakia, N., "Social behavior and brand devotion among iPhone innovators," *International Journal of Information Management*, 30, 6 (2010), pp. 470-480.

[2] Benali-Khoudja, M., Hafez, M., and Kheddar, A., "VITAL: An electromagnetic integrated tactile display," *Displays*, 28, 3 (2007), pp. 113-144.

[3] Blignaut, P. J., "Computerized Self-Administered Questionnaires on Touchscreen Kiosks: Do They Tell the Truth?," in the proceedings of CHI 2004, 2004, pp. 1183-1186.

[4] Bonner, M. N., Brudvik, J. T., Abowd, G. D., and Edwards, W. K., "No-Look Notes: Accessible Eyes-Free Multi-touch Text Entry," *Lecture Notes in Computer Science*, 6030, (2010), pp. 409-426.

[5] Chouvardas, V. G., Miliou, A. N., and Hatalis, M. K., "Tactile displays: Overview and recent advances," *Displays*, 29, 3 (2008), pp. 185-194.

[6] Costlow, T., "Computer Kiosk Expedites Voter Registration," *IEEE Spectrum*, 39, 10 (2002), pp. 26.

[7] Fredkin, E., "Trie Memory," *CACM*, 3, 9 (1960), pp. 490-499.

[8] Guo, S., Falaki, N. H., Oliver, E. A., Ur Rahman, S., Seth, A., Zaharia, M. A., and Keshav, S., "Very Low-Cost Internet Access Using KioskNet," *ACM SIGCOMM Computer Communication Review*, 37, 5 (2007), pp. 95-100.

[9] Hagen, S. and Sandnes, F. E., "Towards Accessible Self-service Kiosks through Intelligent User Interfaces," *Personal and Ubiquitous Computing*, 14, 8 (2010), pp. 715-721.

[10] Janabi-Sharifi, F., Hayward, V., and Wang, Q. Y., "Design and implementation of a graphic-haptic display system," *Displays*, 28, 3 (2007), pp. 118-128.

[11] Johnston, M. and Bangalore, S., "MATCHKiosk: A Multimodal Interactive City Guide," in the proceedings of Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, 2001, pp. 376-383.

[12] Kane, S. K., Bigham, J. P., and Wobbrock, J. O., "Slide Rule: Making Mobile Touch Screens Accessible to Blind People Using Multi-Touch Interaction Techniques," in the proceedings of ASSETS'08, 2008, pp. 73-80.

[13] Kulyukin, V. and Kutiyanawala, A., "Eyes-Free Barcode Localization and Decoding for Visually Impaired Mobile Phone Users.," in the proceedings of International Conference on Image Processing,Computer Vision, and Pattern Recognition., 2010.

[14] Kurtenbach, G. and Buxton, W., "User Learning and Performance with Marking Menus," in the proceedings of CHI'94, 1994, pp. 258-264.

[15] Lee, A., Malik, A., Thai, C., and Manduchi, R., "Universal real-time navigational assistance (URNA): an urban bluetooth beacon for the blind," in the proceedings of

International Conference On Mobile Systems, Applications And Services, 2007, pp. 83-88.

[16] MacKenzie, I. S., "Mobile Text entry using three keys," in the proceedings of NordCHI'02, 2002, pp. 27-34.

[17] McGookin, D., Brewster, S., and Jiang, W., "Investigating touchscreen accessibility for people with visual impairments," in the proceedings of NordiCHI 2008, 2008, pp. 298-307.

[18] Norman, D. A., *The Design of Everyday Things*: The MIT Press, 1998.

[19] O'Neill, E., Kaenampornpan, M., Kostakos, V., Warr, A., and Woodgate, D., "Can we do without GUIs? Gesture and speech interaction with a patient information system," *Personal and Ubiquitous Computing*, 10, 5 (2006), pp. 269-283.

[20] Paradi, J. C. and Ghazarian-Rock, A., "A Framework to Evaluate Video Banking Kiosks," *Omega, Int. J. Mgmt Sci.*, 26, 4 (1998), pp. 523-539.

[21] Park, T. Y., Lee, M. Y., Kim, J. M., and Ha, Y. H., "Design and Implementation of Digital Photo Kiosk System with Auto Color-Correction Module," *IEEE Transactions on Consumer Electronics*, 51, 4 (2005), pp. 1067-1073.

[22] Rana, M. M. and Rana, U., "Accessibility Evaluation of iPhone's User Interface for Visually Impaired," in the proceedings of Internet and Multimedia Systems and Applications - 2009, 2009, pp. 164-167.

[23] Saffer, D., *Designing Gestural Interfaces*: O'Reilly, 2008.

[24] Sánchez, J. and Aguayo, F., "Mobile Messenger for the Blind," *Universal Access in Ambient Intelligence Environments, Lecture Notes in Computer Science*, 4397, (2007), pp. 369-385.

[25] Sánchez, J. and Maureira, E., "Subway Mobility Assistance Tools for Blind Users," in *Universal Access in Ambient Intelligence, Lecture Notes on Computer Science*, vol. 4397, C. Stephanidis and M. Pieper, Eds.: Springer, 2007, pp. 386-404.

[26] Sandnes, F. E., "Directional bias in scrolling tasks: A study of users' scrolling behaviour using a mobile text-entry strategy," *Behaviour & Information Technology*, (2008

[27] Sandnes, F. E. and Aubert, A., "Bimanual Text Entry using Game Controllers: Relying on Users' Spatial Familiarity with QWERTY," *Interacting with Computers*, 19, 2 (2007), pp. 140-150.

[28] Sandnes, F. E. and Huang, Y.-P., "Chording with Spatial Mnemonics: Automatic Error Cor-rection for Eyes-Free Text Entry," *Journal of Information Science and Engineering*, 22, 5 (2006), pp. 1015-1031.

[29] Sandnes, F. E. and Huang, Y.-P., "From Smart Light Dimmers to the IPOD: Text-Input with Circular Gestures on Wheel-Controlled Devices," *International Journal of Smart Home*, 1, 2 (2007), pp. 97-108.

[30] Sandnes, F. E., Jian, H.-L., Huang, Y.-P., and Huang, Y.-M., "User Interface Design for Public Kiosks: An Evaluation of the Taiwan High Speed Rail Ticket Vending Machine," *Journal of Information Science and Engineering*, 26, 1 (2010), pp. 307-321.

[31] Savidis, A. and Stephanidis, C., "Developing dual user interfaces for integrating blind and sighted users: the HOMER UIMS," in the proceedings of The SIGCHI conference on Human factors in computing systems, 1995, pp. 106-113.

[32] Schreder, G., Siebenhandl, K., Mayr, E., and Smuc, M., "The ticket machine challenge? Social inclusion by barrier-free ticket vending machines," in the proceedings of The Good the Bad and the Challenging Conference, 2009.

[33] Slack, F. and Rowley, J., "Kiosks 21: a new role for information kiosks?," *International Journal of Information Management*, 22, (2002), pp. 67-83.

[34] Slay, H., Wentworth, P., and Locke, J., "BingBee, an Information Kiosk for Social Enablement in Marginalized Communities," in the proceedings of SAICSIT 2006, 2006, pp. 107-116.

[35] Thimbleby, H., *Press On: Principles of Interactive Programming*: MIT Press, 2007.

[36] Tinwala, H. and MacKenzie, I. S., "Eyes-free text entry on a touchscreen phone," in the proceedings of 2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH), 2009, pp. 83-88.

[37] Tung, L. L. and Tan, J. H., "A model for the classification of information kiosks in Singapore," *International Journal of Information Management*, 18, 4 (1998), pp. 255-264.

[38] Vanderheiden, G. C., "Use of audio-haptic interface techniques to allow nonvisual access to touchscreen appliances," in the proceedings of Human Factors and Ergonomics Society Annual Meeting Proceedings, 1996, pp. 1266.

[39] WHO, "Consultation on development of standards for characterization of visual loss and visual functioning," World Health Organization 2003.

[40] Wobbrock, J. O., Wilson, A. D., and Li, Y., "Gestures without Libraries, Toolkits or Training: A $1 Recognizer for User Interface Prototypes," in the proceedings of UIST'07, 2007, pp. 159-168.

[41] Yfantidis, G. and Evreinov, G., "Adaptive blind interaction technique for touchscreens," *Universal Access in the Information Society*, 4, 4 (2006), pp. 328-337.

[42] Zhao, S., Dragicevic, P., Chignell, M., Balakrishnan, R., and Baudisch, P., "Earpod: eyes-free menu selection using touch input and reactive audio feedback," in the proceedings of CHI '07, 2007, pp. 1395-1404.