Towards Accessible Self-service Kiosks through Intelligent User Interfaces

Simen Hagen and Frode Eika Sandnes

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

+47 22 45 32 49

+47 22 45 32 05

simen.hagen@iu.hio.no and Frode-Eika.Sandnes@iu.hio.no

http://www.iu.hio.no/~simenhag/, http://www.iu.hio.no/~frodes/

Public self-service kiosks provide key services such as ticket sales, airport check-in and general information. Such kiosks must be universally designed to be used by society at large, irrespective of the individual users' physical and cognitive abilities, level of education and familiarity with the system. The noble goal of universal accessibility is hard to achieve. This study reports experiences with a universally designed kiosk prototype based on a multimodal intelligent user interface that adapts to the user's physical characteristics. The user interacts with the system via a tall rectangular touch sensitive display where the interaction area is adjusted to fit the user's height. A digital camera is used to measure the user's approximate reading distance from the display such that the text size can be adjusted accordingly. The user's touch target accuracy is measured and the target sizes are increased for users with motor difficulties. A Byzantine visualization technique is employed to exploit unused and unreachable screen real estate to provide the user with additional visual cues. The techniques explored in this study have potential for most public self-service kiosks.

Keywords: Self-service kiosk, touch display, intelligent user interface, user aware system, universal design, Byzantine projection

1. Introduction

Advances in technology, reduced hardware costs and increased cost of manual labor have motivated the deployment of self-service kiosks in public spaces. Kiosks are becoming ubiquitous, and most people use them without even thinking about it. Self-service kiosks offer a wide range of services including ticket sales [1], self-service banking [2], photo manipulation and printing [3], interactive city guides [4], low-cost public Internet access [5], assimilation of votes [6], user testing of online books [7], interactive technical manuals [8], collecting responses to questionnaires [9], raising the education level of under-privileged children [10] and education and promotion of child safety [11].

Current kiosks commonly rely on interaction through touch sensitive displays [12]. These displays have until recently been small, and are usually installed with wheelchair users in mind, as required by recommendations, standards, or legislature regarding universally designed physical environments. The requirements for universal design are rooted in the large diversity in physical and cognitive abilities among target users. The largest group of users with reduced function is the elderly, as physical and cognitive abilities weaken with age. Children make up another large group of users with special needs, as their physical and cognitive skills are not fully developed.

Current self-service kiosks are associated with several problems. First, the current practice of a best fit for all regarding the vertical position of the display is not ideal. A kiosk configured for wheelchair users discriminates against tall people – especially tall people with reduced vision. These users need to bend down into an uncomfortable and un-ergonomic posture in order to view the display. Second, the displays may be hard to read for visually impaired users if the text and user interface controls are too small, or have insufficient contrast [13]. On the other hand, if the text is too large, it may be uncomfortable and inefficient to read for customers with 20/20 vision[1]. Moreover, some users with 20/20 vision may feel embarrassed to use an interface clearly designed for visually impaired users. The large print is also infringing on the privacy of the user, as the text may be readable by customers queuing behind. Third, users with motor problems such as Parkinson’s disease may have problems hitting small targets on a touch display. Moreover, users with Parkinson may unintentionally touch the display because of tremors (uncontrollable muscle movements, often affecting the hands, arms or legs). Fourth, due to visual perspective distortion, information displayed above a

---

[1] The term *20/20 vision* is used to refer to a person with normal vision. It is the visual acuity needed to visually separate 2 points with 1 arc minute distance, or 1/16 of an inch, at 20 feet. The numerator is a subjects performance, and the denominator is the norm.

wheelchair user may be hard to read, even if the text is large. Similarly, information displayed below the natural viewing position of tall users may also be hard to read for the same reasons.

This study reports on the experimental design of a self-service kiosk prototype that attempted to address these issues. First, height variations are catered for by a tall rectangular touch sensitive display configured into a portrait position, where the approximate height of use is estimated from the initial point of touch. Then, key parts of the interaction components can be presented at an appropriate height. Moreover, the changes in the distance between the customer and the display were used to adjust the text size as the text was increased when users leaned forward to read. Next, the users' ability to hit the targets accurately was estimated, and the target sizes were adjusted accordingly. Finally, unused display real estate was used to provide additional visual cues by the means of a Byzantine, or inverse, perspective mapping, for improved legibility. Inverse perspective is a visualization method where objects are larger, the further away they are. Lines in the drawing diverge at the horizon, rather than converge as with normal linear perspective [14, 15].

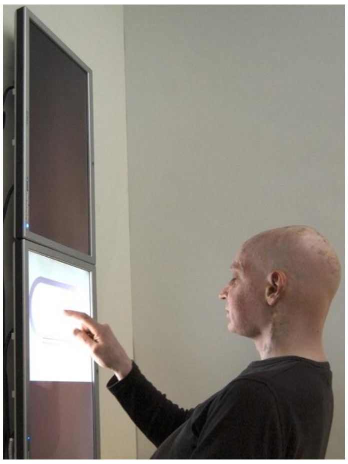

Fig. 1 The interface is positioned at the point of initial touch, in this case at an appropriate height for the user

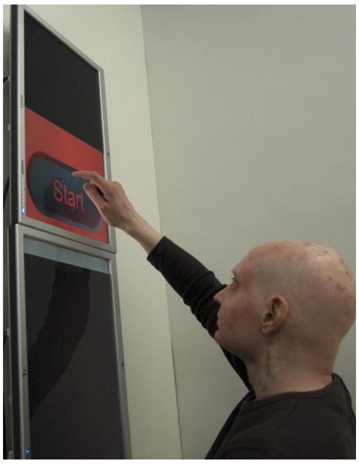

Fig. 2 The interface is also suitable for tall users, here simulated by having an initial touch relatively close to the top of the display

This paper is organized as follows: Section 1 presents the problem and gives a short survey of the literature. In section 2, the challenge of users of different height operating a kiosk is discussed. Section 3 and 4 present our solutions to some of the challenges pertaining publicly available self-service kiosks. Section 5 explains how Byzantine projection is used to make a larger area of the display legible. Section 6 concludes.

2. User height

Current self-service kiosks are equipped with moderately sized touch sensitive displays mounted at fixed heights. The trend is to mount the displays such that they can be reached by wheelchair users who would not be able to reach touch sensitive displays mounted above a certain height (see Fig. 1). Floor selection buttons in lifts are frequently mounted in a similar manner. Unfortunately, such configurations may be uncomfortable for tall users. In an ergonomic optimal scenario, the display would be adjusted according to the user’s vertical level of access. Obviously, mechanical adjustments of the display height are impractical, expensive and time-consuming, being it manual or automatic. Moreover, a mechanical mechanism may be subject to vandalism if the kiosk is placed in a public location without surveillance. Virtual and instant display height adjustments are proposed as a simple remedy to this problem. The solution comprises a long rectangular touch display that is viewable by both wheelchair users (see Fig. 1) and tall users (see Fig. 2), spanning a vertical range of about 70 cm to about 180 cm above the ground. As the user approaches the self-service kiosk, its visual appearance invites the user to touch the display. As the user touches the display, the kiosk immediately detects the touch and centers the user interface on the initial point of touch as shown in Fig. 1 and 2.

One challenge with this approach is how to entice touching of the display in the first instance. In our prototype, we used a combination of text in various sizes, colors and positions, a picture of touch, as well as hand icons to provoke the association of touch among the users approaching the kiosk at a distance (see Fig. 4 a and b). Moreover, the entire display was activated to clearly signal the touchable area of the kiosk. The prototype employed a tiled display comprising two display units [16]. Ultimately, a large tall-screen display in portrait configuration (vertical widescreen), with a continuous surface, could be used.

3. Text size

Unlike a personal computing environment where users are able to configure text size and resolution to fit their personal needs, it is not feasible to personalize a public kiosk intended for anonymous use. Self-service kiosks are usually targeted at users with perfect 20/20 vision, and the user interface designers often add an error margin, i.e., extra large text, such that the kiosk can be accessed by users with less than perfect vision. Reduced vision is a common challenge, as most people’s eyesight is reduced with age. Attention to information display size is especially relevant as most user interfaces are visual.

A clerk will usually be able to recognize that a customer has reduced vision, and assist that customer with deciphering small print on forms, highlight important points, read text aloud, etc. The aim of our self-service kiosk was to mimic the dynamic behavior that is characteristic of good customer service. A digital camera discretely mounted on top of the self-service kiosk observes the user and draws conclusions about changes in reading distance. It is assumed that if the customer is unable to read the information on the display, the customer will move the head closer towards, or further away from, the display. A response to such a motion is to increase the size of text and other elements on the display, allowing the customer to return to a natural distance from the display.

a) Images captured by the video camera

b) Binarized images

c) Vertical projections

Fig 3 Detecting changes in viewing distance. Top row: the user reads the display at normal distance (distance = 148.9 pixels). Middle row: the user reads the display at closer distance (distance = 117.6 pixels). Bottom row: the user reads the display at a close distance (distance = 89.9 pixels)

3.1. Detecting changes in viewing distance

Changes in distance between the display and the user are detected as follows. The image capture starts once the screen is touched. Next, the image is binarized such that the person (the foreground) resumes pixels with ones and the background pixels with zeroes. The floor in our lab had a light color and provided a strong contrast in relation to the user in the foreground. Therefore, binarization was achieved by first inverting the images, and then converting them to gray scale. A Gaussian blur filter was then applied, before finally the images were run through a gray scale filter.

In more challenging environments where there for instance is a low contrast between the floor and the user, the captured image can be compared with a reference image to separate background from foreground and a RGB-HSV color space transformation could be used to further enhance the foreground-background separation. Changes in the background, e.g. people moving around or in the queue behind the user may also provide a challenge, and in a real environment, these elements should be removed.

Next, based on the binarized difference image *I* the vertical projections [17, 18] *V* are computed by summing the pixels of each row, namely

$$V(y) = \sum_x I(x,y)$$

Then, the mean of the projections provides a rough representation of the central position of the user's head in relation to the kiosk in pixels, namely

$$\overline{V} = \frac{\sum_y V(y) \cdot y}{\sum_y V(y)}$$

The head position calculations are illustrated in Fig. 3, which shows a user at a normal distance from the display (top), leaning forward towards the display (middle) and very close to the display (bottom). The example also shows the binarized images and the corresponding vertical projections. The distances computed from the projections in the example using the technique are 148.9, 117.6 and 89.9 pixels, respectively.

The head distance is measured at regular intervals, in our case every 500 ms, and the mean head distance d_i is calculated for the duration of which view *i* is displayed. In order to determine a suitable text size for the next view, the mean display distance for the last view d_i is then compared to the mean display distance for the one before d_{i-1} such that the text is scaled using the scaling factor *t(d)*

$$t(d) = \begin{cases} 0.5625, & d_i > (1+k)d_{i-1} \\ 1.0000, & (1-k)d_{i-1} \le d_i \le (1-k)d_{i-1} \\ 1.7777, & d < (1-k)d_{i-1} \end{cases}$$

where *k* is a constant and set to 0.2. Moreover, the text size is not altered if the standard deviation of the distance measurements for the view exceeds 30% of the distance in order to avoid noise caused by users moving around, talking to friends, look for items in their bags, etc. The reduction factor 9/16 and expansion factor 16/9 were found through experimentation and allows 200-point text to be reduced to 10-point text in 6 steps.

a) Welcoming view

b) Initial view with a text size of 200 points

c) Initial view with a lower center of interaction

d) The first view with a text size of 100 points.

e) The second view with a text size of 112 points.

Fig 4 Dynamically altering the text size and center of interaction.

3.2. Text size adjustments

Most self-service transactions should be quick and easy to execute, and involve only a minimal number of steps. The initial view is set up with a text size that is larger than what is necessary for most users, in order to include users with reduced vision. If a particular view has insufficient space to display all the information in the active interface area, the text size is reduced (see Fig. 4). As the user progresses though the views, the display distance is monitored. If the user is consistently leaning forward during a view, it is assumed that the text is hard to read and that the lower text size limit for the user is reached. Consequently, the previous, and larger, text size is used for the remaining steps.

Some visually impaired users employ screen magnifiers typically with a magnification factor of eight. If we assume that a user with 20/20 vision reads text with a physical height of 1 cm comfortably at an arm lengths distance from the screen, then this text should be 8 cm for these visually impaired users. The welcoming text on the display should therefore be at least 8 cm tall when rendered on the display (the actual typographic size will vary according to display dimensions and resolution used (PPI or pixel per inch), and the choice of font and font size).

It is necessary to be within a distance of approximately an arm’s length to operate a self-service kiosk. The length of an arm is proportional to the height of the person, and therefore the distance to the screen may be related to the height of a person. For office work, the distance from the display to the users’ eyes is recommended to be 60 to 80 cm [19], which is approximately an arm’s length. Touch interaction require the user to be somewhat closer.

Presbyopia users, who typically have problems focusing on things that are close, may move away from the display in order to properly see the screen. The system has a minimum font size that should be legible even if the user is standing a few steps back.

Finally, when altering text size dynamically it is important to conserve an overall consistent visual structure of the views, which helps the users navigate and recognize views.

4. Touch target size and accuracy

Many forms of motor impairments exist. This study focuses on motor impairments that result in reduced target hitting accuracy. For instance, users with

Parkinson's disease are characterized as having "shaking hands" (tremors), and are therefore unable to hit small targets [20]. Smaller children are also known to have uncontrolled movements and may find it more difficult to hit small targets [21]. Moreover, self-service kiosks in public spaces are often used by users under stress with a high heart rate. The targeting accuracy of a user is reduced with increasing heart rate and any stressed person may find it hard to hit small targets, especially if they are in a hurry. Fitts' law describes the tradeoff between accuracy and speed with rapid targeting motions relevant to touch [22].

One aim of our self-service kiosk was to adapt to users with reduced targeting ability. Again, the principle of increasing information content and decreasing target sizes was employed, as there is no room to conduct explicit testing of a user's targeting accuracy, and the user may only need to conduct a handful of touches in order to complete a transaction.

The procedure works as follows. For each step, the accuracy of the touch is calculated. If the accuracy is good, the target size can be reduced. The target size remains at the same level if the accuracy is just acceptable, and if the target accuracy is unacceptable, the view is displayed again with larger targets. The accuracy *A(p)* of a touch *p=(x,y)* was defined as

$$A(p) = \frac{|p - p_0|}{R(p)}$$

where p_0 is the center of the target (x_0,y_0), and *R(p)* is the distance from the center of the target to the touching point *p*. The text was then scaled by the scaling factor *s(p)* defined as

$$\mathrm{s(p)} = \begin{cases} 0.5625, & A(p) < 0.5 \\ 1.0000, & 1.0 \geq A(p) \geq 0.5 \\ 1.7777, & A(p) > 1.0 \end{cases}$$

Hence, if the user hits a point closer to the center than the border, the target size is reduced. If the user hits closer to the border of the target, the target size remains unchanged. If the user hits outside the border, the target size is increased.

In order to avoid erroneous clicks on nearby objects, all objects should be surrounded by sufficient space. The amount of space is dynamically adjusted in unison with the size of the target, using the same scaling factor as for the targets.

a) The display observed from a tall user's perspective. The perspective causes the text to be visually distorted, and harder to read than if seen from a perpendicular angle.

b) The same display observed from a tall user's perspective, after Byzantine perspective correction is applied. The text is easier to read than in figure 5(a).

Fig 5 Displayed text with and without inverse perspective correction

5. Viewing unreachable display real-estate

Most of the available real estate on the display will remain unused if only the area immediately adjacent to the user's face is utilized. The upper parts of the display remain unused for short users and the lower part for tall users. Clearly, the parts of the display further away from the viewer are harder to read because of the text size to distance ratio. Moreover, the perspective view of these far-away display areas will transform the image, making it hard to decipher.

However, by magnifying the parts farthest away to compensate for the distance, and use inverse, or Byzantine, perspective mapping [14, 15] to compensate for the projection distortion, these previously unused display areas will become exploitable as shown in Fig. 5. The inverse projection counterbalances a disappearing viewing plane such that the area to be viewed appears perpendicular and opposite to the viewer. It has been shown that, depending on the angle, the legibility of the text is improved by up to 60% when the text is viewed from the side using an inverse projection mapping [15].

In particular, the extra area can be used to draw attention to other parts of the kiosk, such as where to insert a credit card or where to collect tickets or printed photographs. Alternatively, the extra area can be used to display information requiring more space, such as maps or transport line charts. Other scenarios include displaying current information such as time, currently departing train, next bus, news headlines, etc. Clearly, such extra display real estate should be used sparingly as the user is mostly able to focus on one thing at any one time.

6. Discussion and future work

We have presented a self-service kiosk prototype that adapts to the users needs. The kiosk analyzes the movement of the users, adapts font sizes and user interface element positions to the needs of the user. Inverse perspective mapping is also used to allow easy reading of interface elements located furthest away from the user.

However, many extensions can be envisaged. The users’ positions were used to estimate their abilities to read the presented information. Instead, or in addition, one could detect the users’ gaze. By knowing where the users are looking, one could estimate whether they are able to find, or read, the desired information. In [23], the authors describe how to help the users when they are looking for information by using a robot’s gaze. However, current gaze technology does not work with all seeing users.

Building audio into the user interface may help users with severe visual impairments and blindness. Audio assistance is currently the focus of several promising developments, and the authors would like to explore this further.

Many other self-service kiosk design challenges remain unresolved. These include the problems associated with the hand obstructing information when using a touch display [24], accidentally and prematurely touching the screen for users with motor problems, and the difficulty blind users experience with touch displays due to the lack of tactile feedback. The method of detecting the user’s distance to the display is simple, and has not been tested in a realistic environment. Challenges such as removing the background and people in the background, must be addressed.

Some of the challenges with achieving universally designed public self-service kiosks were discussed. Strategies for on-the-fly user interface configuration based on the users' height, reading distance, and ability to hit a target accurately were explored. The current study is based on low-fidelity prototypes, and more work is needed to further determine the practicality of these techniques.

References

1. Sandnes, F.E., et al., *User Interface Design for Public Kiosks: An Evaluation of the Taiwan High Speed Rail Ticket Vending Machine.* Journal of Information Science and Engineering, 2010. **26**(1).
2. Paradi, J.C. and A. Ghazarian-Rock, *A Framework to Evaluate Video Banking Kiosks.* Omega, Int. J. Mgmt Sci., 1998. **26**(4): p. 523-539.
3. Park, T.Y., et al., *Design and Implementation of Digital Photo Kiosk System with Auto Color-Correction Module.* IEEE Transactions on Consumer Electronics, 2005. **51**(4): p. 1067-1073.
4. Johnston, M. and S. Bangalore. *MATCHKiosk: A Multimodal Interactive City Guide*. in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*. 2001: ACM Press.
5. Guo, S., et al., *Very Low-Cost Internet Access Using KioskNet.* ACM SIGCOMM Computer Communication Review, 2007. **37**(5): p. 95-100.
6. Costlow, T., *Computer Kiosk Expedites Voter Registration.* IEEE Spectrum, 2002. **39**(10): p. 26.
7. Schokz, J. *Kiosk-Based User Testing of Online Books*. in *Proceedings of the 16th annual international conference on Computer documentation*. 1998: ACM Press.
8. Greenough, R.M. and B. Tjahjono, *An interactive electronic technical manual for an advanced aerospace assembly machine.* Int. J. Adv. Manuf. Technol., 2007. **33**(9-10): p. 1045-1055.
9. Blignaut, P.J. *Computerized Self-Administered Questionnaires on Touchscreen Kiosks: Do They Tell the Truth?* in *CHI 2004*. 2004: ACM Press.
10. Slay, H., P. Wentworth, and J. Locke. *BingBee, an Information Kiosk for Social Enablement in Marginalized Communities*. in *Proceedings of SAICSIT 2006*. 2006: SAICSIT.
11. Gielen, A.C., et al., *Using a Computer Kiosk to Promote Child Safety: Results of a Randomized, Controlled Trial in an Urban Pediatric Emergency Department.* Pediatrics, 2007. **120**: p. 330-339.
12. Saffer, D., *Designing Gestural Interfaces*. 2008: O'Reilly.
13. Lin, C.-C., *Effects of contrast ratio and text color on visual performance with TFT-LCD.* International Journal of Industrial Ergonomics, 2003. **31**(2): p. 65-72.

14. Nacenta, M.A., et al. *E-conic: a perspective-aware interface for multi-display environments*. in *Symposium on User Interface Software and Technology*. 2007: ACM Press.
15. Larson, K., et al. *Text in 3D: some legibility results*. in *Conference on Human Factors in Computing Systems*. 2000: ACM Press.
16. Li, K., et al., *Building and Using A Scalable Display Wall System.* IEEE Computer Graphics and Applications, 2000. **20**(4): p. 29-37.
17. Huang, Y.-P., et al., *An Intelligent Strategy for Checking the Annual Inspection Status of Motorcycles based on License Plate Recognition.* Expert Systems With Applications, 2009. **36**(5): p. 9260-6267.
18. Huang, Y.-P., et al., *Back Propagation based Real-Time License Plate Recognition System.* International Journal of Pattern Recognition and Artificial Intelligence, 2008. **22**(2): p. 233-251.
19. Anshel, J., *Visual ergonomics in the workplace*. 1998: CRC.
20. Keates, S. and S. Trewin. *Effect of age and Parkinson's disease on cursor positioning using a mouse*. in *ACM SIGACCESS*. 2005: ACM Press.
21. Hourcade, J.P., et al., *Differences in pointing task performance between preschool children and adults using mice.* ACM Transactions on Computer-Human Interaction, 2004. **11**(4): p. 357 - 386.
22. Fitts, P.M., *The information capacity of the human motor system in controlling amplitude of movement.* Journal of Experimental Psychology, 1954. **47**: p. 381-391.
23. Yonezawa, T., et al. *GazeRoboard: Gaze-communicative guide system in daily life on stuffed-toy robot with interactive display board*. 2008.
24. Wigdor, D., et al. *LucidTouch: A See-Through Mobile Device*. in *UIST 2007*. 2007: ACM press.