1版

様　式　C－１９、F－１９、Z－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成 ２８ 年　 ６ 月 ３０ 日現在

機関番号： ２６４０２
研究種目： 基盤研究(C) (一般)
研究期間： 2013～2015
課題番号： ２５３３０２４１
研究課題名（和文） Assisting Blind People to Interact with Public Displays

研究課題名（英文） Assisting Blind People to Interact with Public Displays

研究代表者
金　起範（KIM, kibum）

高知工科大学・総合研究所・客員研究員

研究者番号：２０６４６９６３

交付決定額（研究期間全体）：（直接経費）　　3,800,000円

研究成果の概要（和文）：大型対話型ディスプレイ・システムは、晴眼者に多大な利益をもたらす一方、視覚障がい者にとって依然として近づき難い。私たちは、視覚障がい者に同様な利益をもたらすことを目指し、壁掛け型の大型ディスプレイ上の目標物に視覚障がい者がアクセスするためのジェスチャー入力と触振動覚フィードバックの実行可能性を調査してきた。
本研究結果は、視覚障がい者が垂直の大型対話型ディスプレイ上で情報を入手する能力を増強する私たちの提案技術の可能性を裏付けるものである。この初期段階の研究が、近い将来、視覚障がい者の大型公共ディスプレイ上の情報入手に晴眼者と同様に役立つ第一歩となることが私たちの願望である。

研究成果の概要（英文）：The rapid growth of large interactive display systems in our daily lives enables sighted people to walk up and derive great advantages spontaneously and naturally. On the other hand, this technology remains largely inaccessible to people with visual impairments. In an effort to deliver the same benefits to people with visual impairments, we have explored the feasibility of gesture input and vibrotactile feedback to help them access targets on a large wall-mounted display through several experiments. Based on the experiment results, we could publish a couple of papers to international journals.

Our results confirm the potential of our proposed techniques to enhance the ability of people with visual impairments to access information on large interactive vertical displays. It is our aspiration that this initial work will serve as the first step towards granting them equal access to information on large public displays in the near future.

研究分野：Human Computer Interaction

キーワード：HCI　CSCW　Assistive Technology　Blind Users　Large Vertical Display　Haptic　Air Gesture　Handheld Device

１．研究開始当初の背景

There exist assistive technologies for people with visual impairments in order to identify cases relevant to large vertical displays. Interactive devices for assisting people with visual impairments have various form factors. Small, portable ones include a white cane (Astler et al., 2011; Fernandes et al., 2011), wearable widgets such as a vibrotactile glove (Krishna et al., 2010) and a wearable camera (Jeon et al., 2012), and handhelds such as a PDA (Ghiani et al., 2008; Sanchez et al., 2008), a smartphone (Azenkot et al., 2011a; Southern et al., 2012), force disk (Amemiya and Sugiyama, 2009) as well as Wii Remote (Morelli et al., 2010). Large displays like tabletops (Kane et al., 2011; Kane et al., 2013b; Manshad et al., 2013) have also been used.

The form factor appears to be correlated to the nature of the task involved. For example, to assist blind users with photo taking and photo sharing, smartphones have become very popular devices. EasySnap (White et al., 2010) and Accessible Photo Album (Harada et al., 2013) were developed for the Apple iPhone by employing iOS's VoiceOver screen reading functionality. PortraitFramer (Jayant et al., 2011) on Android phones explored the usefulness of haptic and audio cues for proper people positioning in group photos. In addition, the embedded camera and smartphone portability made them useful as mainstream crowdsourcing devices for collecting information regarding everyday visual challenges faced by people with visual impairments (Bigham et al., 2010; Brady et al., 2013).

On the other hand, large tabletop displays were used for collaborative learning or target acquisition by people with visual impairments. For example, TIMMs (Trackable Interactive Multimodal Manipulatives) provides multimodal feedback (e.g., speech, sound/music, vibration/force feedback) to enable collaborative learning between sighted and blind students (Manshad et al., 2013). Access Overlays (Kane et al., 2011) and Touchplates (Kane et al., 2013b) support easy target selection on a tabletop, e.g., selecting a location on a map, by providing speech or tactile feedback.

In most of the assistive technologies we reviewed, the auditory and/or haptic channels are the de-facto channels for providing sensory feedback. Audio feedback, such as text-to-speech, is quite effective for communication between the system and users with visual impairments. For example, Blobby (Nicolau et al., 2009) uses familiar and easily understandable speech, behaving like a blind companion, to help people with visual impairments navigate through unfamiliar places. Auditory icons, or earcons, are also useful. The SoundsRight CAPTCHA (Completely Automated Public Turing tests to tell Computers and Human Apart) uses earcons taken from environmental sounds or sounds related to a specific concept (e.g., a train, rain, animals, etc.). Users with visual impairments were able to achieve success rates higher than 90% using this technology (Lazar et al., 2012).

Audio feedback has often been augmented by haptic feedback. It was shown that people with visual impairments could perceive an audio-based representation of a bar graph using a pointing device if provided with adequate tactile feedback (Wall and Brewster, 2006). Haptic feedback presented by a tactile array can enable people with visual impairments to perceive graphical information and form a mental model for visual imagery. For example, GraVVITAS demonstrated that the combination of a touch sensitive tablet and a data glove with vibrating actuators could be an effective technique for representing tactile graphics to users with visual impairments (Goncu and Marriott, 2011).

There has been a body of research pertaining to eyes-free target acquisition interaction (Bahram et al., 2012; Cockburn et al., 2011; Fiannaca et al., 2013; Folmer and Morelli, 2012; Li et al., 2009, 2010; Manduchi and Coughlan, 2014). CAVIAR (Bahram et al., 2012) supported acquiring objects in the peripersonal space. In CAVIAR, a wristband with its vibrotactile actuators generates continuous stimuli to guide the user's hand. The wristband was directed via Bluetooth from a mobile phone which recognized and tracked the hand and objects using computer vision. Air pointing (Cockburn et al., 2011) enables eyes-free interactions, where users rely on proprioception and kinesthesia rather than vision. They gradually reduced the amount of visual feedback until there was none. The effect of different feedback techniques on accuracy and learnability was suggested as the extension of their study. Fiannaca et al. (2013) and Folmer and Morelli (2012) presented haptic feedback to point out the location of a virtual object in tactile-proprioceptive displays and evaluated multilinear target-scanning in a plane in front of the user. Their vibrotactor used pulse delay and frequency to provide directional vibrotactile feedback. Virtual Shelves (Li et al. 2009, 2010) was for selecting nonvisual objects by positioning a Wiimote within a virtual circular hemisphere defined in front of the user. On the other hand, the Last Meter (Manduchi and Coughlan, 2014) investigated the effects of frame rate and camera field of view on the ability of users with visual impairments to search for visual targets.

２．研究の目的

With considerable progress in display technologies and interaction techniques, we are observing increasing affordability and availability of large interactive displays in our everyday lives. The rapid growth of such large public displays allows us to access a variety of information in diverse places and the contents are now much more interactive. However, the proliferation of large interactive displays also creates great challenges for people with visual impairments, who require equal access to information on displays that are predominately visual. Accommodating the special needs of people with visual impairments is not only socially valuable but also produces more effective and widely useful interfaces for everyone (ACM code of ethics and professional conduct, 2015).

Assistive technology research has come a long way and has yielded many effective interfaces. Nevertheless, assistive technology for *large vertical displays* is still sparse and poorly supported in part due to the lack of a good understanding of the challenges faced by *people with visual impairments*. This project investigates how to facilitate a target-aiming task on a large wall-mounted display by people with visual impairments. The ability to correctly select a target on an interface is the first step toward further manipulation and it is fundamental in most modern graphical user interfaces. Specifically, we present three target-aiming techniques – Random, Cruciform, and Radial – using natural gesture input aided by directional vibrotactile feedback. In all three techniques, our users with visual impairments point their hand at the large wall-mounted display, and this mid-air, non-contact gesture is tracked by a computer vision system and mapped to a cursor position onto the screen like a mouse using ray-casting. To determine the best search direction, directional vibrotactile stimuli are delivered by means of a mobile phone held in the pointing hand. The three target-aiming techniques differ in the geometric path along which the search is performed and the way vibrotactile feedback is provided for guidance.

Our lightweight, low-cost interface allows users with visual impairments to dynamically access targets on a large vertical display in a 3D environment. This was confirmed by an experiment in which the three haptic target-aiming techniques were compared and by another experiment in which speech and haptic feedback on the target-aiming technique were compared.

This project was, to the best of our knowledge, among the few of its kind in that it’s concerned with the use of the large vertical display by people with visual impairments. The findings of this project will serve as an initial step towards enhancing the interactivity of large vertical displays for people with visual impairments, with an ultimate goal of granting them equal access to public displays that are currently accessible almost exclusively to sighted people.

３．研究の方法

In our target-aiming techniques, guidance cues are provided by means of vibrotactile stimuli produced on a mobile device held in the pointing hand. An intuitive approach to delivering such directional guidance is spatial coding, wherein a movement direction is represented by the location of vibrotactile stimulation or its positional changes over time. Humans have well-defined innate mapping between the stimulated body site and its corresponding egocentric orientation (Choi and Kuchenbecker, 2013). However, since mobile phones are made of rigid materials and vibration is propagated along their surface, using localized stimulation sites for directional cues is not feasible. An alternative is to rely on *vibrotactile flows*, which refer to vibrotactile sensations moving from one end of the device to the other (Seo and Choi, 2010). In particular, vibrotactile flows can be generated by only two actuators. Our design uses four actuators to represent four different directional cues. Vibrotactile flows can be generated in four directions, upward or downward using the top and bottom actuators, and leftward or rightward using the left and right actuators.

Vibrotactile flows in each direction are rendered using the two actuators at the corresponding starting and ending positions. For this purpose, any of three methods can be used: amplitude inhibition (Seo and Choi, 2010), time inhibition (Kim and Kim, 2012), and amplitude inhibition with frequency sweep (Kang et al., 2012). We use amplitude inhibition because it is simple and its perceptual effects are relatively well known (Seo and Choi, 2013). In this method, the intensity of an actuator at the starting position is monotonically decreased and the intensity of an actuator at the ending position is monotonically increased, using the following synthesis equations (Seo and Choi, 2013):

$$a_1(t) = a_{max}\left(\frac{t}{T}\right)^{\gamma}.$$

$$a_2(t) = a_{max}\left(1-\frac{t}{T}\right)^{\gamma}$$

where $a_1(t)$ and $a_2(t)$ are the respective accelerations of two actuators at time $t \in [0, T]$. a_{max} is the maximum acceleration, and T is the flow duration. Here, $a_1(t)$ is monotonically increased as $a_2(t)$ is monotonically decreased. This equation results in a vibrotactile flow linearly moving from the position of actuator 2 to that of actuator 1. The γ parameter determines the increasing/decreasing rates of the two vibrations and also has an effect on the overall perceptual attributes of vibrotactile flow, such as perceived travel distance and the confidence of flow-like sensation (Kang et al., 2012). After several pilot tests with blind users, we decided to use $\gamma = 0.65$, which allows for the perception of a smooth directional movement. For more robust identification, we provide an identical flow twice with a short inter-stimulus interval (0.3 s). This consists of a vibrotactile stimulus for one guidance cue.

When the user changes the movement quickly, a vibrotactile cue can be initiated even if the play of the preceding cue has not yet finished. If this occurs, we stop the preceding cue immediately, and then play the following cue. To help the users distinguish the change, a short strong vibration (0.1 s, a_{max}) is provided between the two cues using the actuator at the top of the mockup. Since the perceived direction of flow depends on hand orientation, the user needs to maintain his/her grasping posture and orientation consistently, and this is a natural behavior observed during the experiment.

We compared three target-aiming techniques to help people with visual impairments aim at a target on a large vertical display. The techniques are Random, Cruciform, and Radial. In all of them, users interact with the display using mid-air gestures while guided by vibrotactile feedback on a mobile device which is held in their pointing hand.

3.1. Random technique

This is a freestyle search in which users move the hand randomly in the air to aim at a target, receiving no guidepost to the target position. Microsoft Kinect's skeletal tracking supplies the coordinates of hand and wrist positions, which are used to calculate a ray-casting vector along the same direction as a projected line from the wrist to the hand. This allows the user's unrestricted hand movement to control the cursor on the display in a congruent manner for target selection. If the hand touches a target, the user receives confirmation feedback by means of a short vibration from the mobile device and a bell tone. The Random technique worked as a baseline for our experiment and was experimentally compared with our two enhanced target-aiming techniques below.

3.2. *Cruciform technique*

With a flat-screen tabletop computer, blind users showed preferences for touch-based input gestures that used screen corners and edges as landmarks (Kane et al., 2011). However, with a large vertical display, access to the edges of the display is not always feasible. Given the users' height and the elevation of the display, users with visual impairments may not be able to easily access the edges of the display. To compensate for people with visual impairments' limited access to the edges of a large vertical display, but still to make use of the benefits of *linearization of the search* (Kane et al., 2011), we devised the Cruciform technique.

The Cruciform technique extends a target location by stretching its x-position horizontally and y-position vertically, making the two vibrotactile feedback areas intersect at the target|. As a result, users do not need to slide their hand along the edges of the display to access the target. Instead, they can move their hand horizontally or vertically from its current position until it reaches the vertical or horizontal extension of the target. When this occurs, the haptic device delivers vibrotactile guidance in the direction of the target from the current hand position, which is leftward, rightward, upward, or downward. Then, the user moves his/her hand accordingly. When the pointing hand reaches the target, s/he feels the same feedback as that of the Random technique for successful target selection.

3.3. *Radial technique*

The Radial technique adopts a compass metaphor in guiding the user's hand towards the target. This technique provides a guidepost from the center of a user's shoulder. The user starts by drawing a circle in the air to find a direction to the target from the center of the shoulder. If the user points along the direction to the target, the mobile device delivers one of two kinds of vibrational feedback (i.e., outward or inward) depending on the current hand position relative to the target position. Inward and outward vibrations are delivered pairing actuators (e.g., top & left actuators and bottom & right actuators) depending on the current hand position relative to the target position. Then the user stretches or contracts the arm accordingly to reach the target. If the hand moves outside of this direction, the stimuli for guiding the user will stop, and when the hand enters in the right direction, the directional vibrotactile feedback is generated again. When the pointing hand reaches the target, the user feels the short vibration and hears the bell tone, as with the other two techniques.

The Cruciform and Radial techniques reduce 2D search dimensionality to two linear searches of a Euclidean coordinating system (x and y) or a polar coordinate system (angle and distance). This mapping allows users to first select a ballpark target area and then to locate it more precisely. The Radial technique supports each user with a customized guidepost towards the target, based on his/her height and standing position in front of the display, while the Cruciform technique provides a general guidepost for all the users regardless of their height or position.

４．研究成果

We evaluated the three target-aiming techniques with 11 participants with visual impairments. The Random technique worked as a baseline against which the other two target-aiming techniques were compared. We also examined whether participants could improve their target aiming in terms of target finding time and failure rate.

4.1. *Target finding time*

Target finding time was defined as the duration from the moment when the pointing cursor left the starting square to the moment the cursor entered the target square. The target-aiming technique had a significant effect on the mean target finding times ($F(2,20) = 15.92$, $p < 0.0005$). The overall mean target-finding times were 30.6s (Standard Error (SE) = 4.0s) for Random, 16.3s (SE = 2.1s) for Cruciform, and 14.9s (SE = 1.8s) for Radial. Post-hoc pair wise comparisons showed that both Cruciform and Radial were significantly faster than Random ($p < 0.01$ and $p < 0.001$, respectively). However, no significant difference was found between Cruciform and Radial.

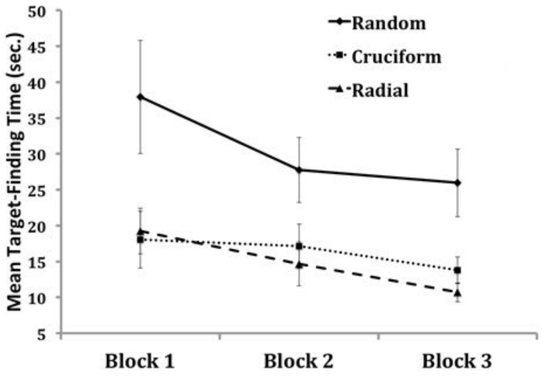

Fig. 1. Mean target-finding times for three target-aiming techniques. Error bars show ±standard error.

Repeating the tasks in three blocks helped the participants improve their target-finding time. However, the p-value approached but did not reach statistical significance ($p = 0.053$). The mean target-finding times were 25.1s (SE = 4.0s) for the first block, 19.9s (SE = 2.0s) for the second block, and 16.8s (SE = 2.0s) for the third block. Figure 5 shows the mean task completion times for the three target-aiming techniques over the three repeated blocks. The Radial technique showed a trend to be faster than Cruciform as participants repeated the tasks, but did not reach statistical significance.

Target size had no significant effect on the results. It took an average of 25.7s (SE = 3.8s) to aim at the small target, 18.1s (SE = 2.2s) to aim at the medium target, and 18.0s (SE = 3.2s) to aim at the large target.

4.2. Failure rate

During all the trials, there was only one case where a participant failed to find the target within the specified time limit (3 mins). This occurred when the participant tried to aim at a small target using the Random technique.

4.3. Subjective preferences

We composed five questions by referring to the ISO9241-400 evaluation questionnaire, such as learnability, easy of use, physical demand, mental demand, and desire to use (ISO, 2007). Results revealed no statistically significant difference in any of the questions among the three techniques. However, participants generally answered favorably in the following order: Radial, Cruciform, and Random.

The rapid growth of large interactive display systems in our daily lives enables sighted people to walk up and derive great advantages spontaneously and naturally. On the other hand, this technology remains largely inaccessible to people with visual impairments. In an effort to deliver the same benefits to people with visual impairments, we have explored the feasibility of gesture input and vibrotactile feedback to help them access targets on a large wall-mounted display. Our results confirm the potential of our proposed techniques to enhance the ability of people with visual impairments to access information on large interactive vertical displays. It is our aspiration that this initial work will serve as the first step towards granting them equal access to information on large public displays in the near future.

５．主な発表論文等

〔雑誌論文〕（計 3 件）

①Kim, K., Ren, X., and Gao, Y. “ShifTable: A Natural Remote Target-Selection Technique on Large Displays”, Interacting with Computers, 28(2), pp. 181-193. March 2016. doi:10.1093/iwc/iwv024. (SCI Journal)

② Kim, K., Ren, X., Choi, S., and Tan, H. “Assisting People with Visual Impairments in Aiming at a Target on a Large Wall-Mounted Display”, International Journal of Human-Computer Studies, Vol. 86, pp. 109-120. February 2016. doi:10.1016/j.ijhcs.2015.10.002 (SCI Journal)

③ Kim, K., and Ren, X. “Assisting Visually Impaired People to Acquire Targets on a Large Wall-Mounted Display”, Journal of Computer Science and Technology, 29(5), pp. 825-836. September 2014. (SCIE Journal)

６．研究組織

(1)研究代表者
金　起範（KIM, kibum）
高知工科大学・総合研究所・客員研究員
研究者番号 ：20646963

(2)研究分担者
任　向実（REN, xiangshi）
高知工科大学・工学部・教授
研究者番号：00287442