ASSISTIVE TECHNOLOGY FOR VISUALLY-IMPAIRED PEOPLE IN ACCESSING GRAPHICAL MATERIAL USING TOUCH-BASED DEVICES

By

HariPrasath Palani

Thesis Advisor: Dr. Nicholas A.Giudice

Abstract

Although critical developments have been made in providing accessible graphics for blind and visually-impaired people, gaining access to graphical material such as maps, graphs and images, has not been widely addressed. The approaches that have addressed the development of accessible graphics have been plagued with shortcomings that have hindered progress in reaching these users. For instance, many of these approaches involve purchase of expensive single-purpose hardware whose design and development was primarily driven by engineering principles rather than theoretical knowledge of human information processing and awareness about the needs and behaviors of end-user's. As the visually-impaired demographic is estimated to number around 12 million people in the U.S. and 285 million people worldwide, the need for developing devices that are both accessible and usable is critical for educational, social, and vocational purposes. This thesis work studies and evaluates an inexpensive, intuitive, and dynamic approach called a vibro-audio interface, for providing non-visual access to graphic material. The interface allows user to freely explore graphical information on the touch screen of a commercially available touch-based devices and synchronously triggering vibration patterns and auditory information, whenever an on-screen visual element was touched. Four studies were conducted with the vibro-audio interface to assess the usability and efficiency of the interface. The first experiment investigated the ability of the interface in conveying graphical information such as shapes, relations, patterns and orientations. The results provide clear evidence that the vibro-audio interface is equivalent in performance when compared with the most accepted mode of non-visual learning, traditional hardcopy braille output. The first experiment was designed in such a way that the graphical materials displayed were not greater than the screen size of the device. This brought up the question of whether the interface would perform in the same efficient way if the graphics are bigger than the screen size, as most of the real world graphics such as shapes, maps, and images will be in varying sizes. In order to address this issue, experiments were conducted to assess the efficacy of the interface in providing map information with different panning and zooming modes. The second experiment focused on different non-visual panning methods and the third experiment focused on non-visual zooming methods. The results suggested that the interface is efficient in conveying map information and can develop an accurate mental representation of the map perceived. The results also signified that two-finger panning mode and button-based zooming mode are more efficient in assisting map transformation for visually-impaired users. The fourth experiment demonstrates the efficacy of the interface in assisting real-world navigation. The users were able to navigate efficiently with the help of the cognitive map that was developed by learning the maps as perceived through the vibro-audio interface. The results also show better performance than conditions using audio-only navigation assistance and conditions without any navigation assistance, suggesting that the vibro-audio interface is a viable multimodal solution for presenting dynamic visual information and supporting accurate spatial learning and the development of flexible mental representations of otherwise inaccessible graphical material.

Introduction

Whilst the majority of our everyday activities rely on some form of digital device to gain information, it is inevitable to accept the fact that the information provided is, in most cases visually based. Accessing such information is always a huge hurdle for visually-impaired users. Part of the issue is resolved by the advent of screen readers like Jaws, and text input software like Dragon. Though these software provide access to text-based output and input, they cannot solve the other part of the issue which is to read the graphical content. The fact that many critical information such as maps, graphs, images are provided in the form of graphics, attributes to the need for developing usable software and hardware to read graphical material. Gaining information to such graphics has been a major hurdle for visually-impaired users. Many research projects have been carried out to resolve this issue, but most of them have focused on engineering product which involves the user to be dependent on an additional gadget. In addition to the fact that these additional gadgets are expensive, they are single purpose and are not portable. The advent of touch-based (smart phones and tablets) devices has opened up new possibilities for resolving the above mentioned issue. The touch-based devices are in-expensive, off-the-shelf, multi-purpose and with its capabilities such as Haptic, Audio, GPS, and other sensors in addition to adopting universal design principles it can be used to solve many accessibility issues. This thesis work focuses on utilizing the capabilities of such touch-based devices to provide solutions for accessibility problems relating to visually-based graphical information faced by visually-impaired users. In particular, the research focus is on providing access to graphical material such as images, graphs, and maps. The idea is to compensate the visual component of the graphical material using other modalities namely Haptic Audio and Kinesthetic. A non-visual interface called Vibro-Audio Interface was developed to accomplish the task. The interface can be readily implemented on any commercially available smart touch devices (e.g. Smart phones and Tablets).

Maps are inevitable for navigation in both indoor and outdoor environments. Maps are complex graphics as they must convey a wide range of information. Providing all the map information in the form of haptic information can be confusing and make it unusable. Though the structure of the indoor and outdoor maps are similar (Yang et al., 2011), this work is focused only on indoor maps. It is necessary to understand the important component of the map and to convey each of the components in either haptic or audio modality. It is also necessary to understand whether a visually-impaired person can use the interface. Empirical research on touch-based interface has addressed these issues(Su et al., 2010, Poppinga et al., 2011, Goncu, 2011) and also focused on design guidelines such as optimal line width (0.35 inch) for providing haptic cues(Raja et al., 2011). However, these studies did not require development of cognitive map to perform spatial tasks. Thus, part of the research work in this thesis is to validate the design guidelines from the empirical study results using the newly developed vibro-audio interface and to evaluate the efficacy of interface in developing accurate spatial representation of the objects perceived.

In order to accomplish this, a human behavioral study was conducted. The scenario was that, *"If a visually-impaired person is enrolled in a geometry class, how the student can follow the class when the discussion is related to a graphic element such as bar graph or geometric shapes?"* There can be multiple solutions for this scenario such as using an audio graph or braille graph. But, these solutions have potential shortcomings such as distracting and not being dynamic and portable, which is why they are not widely used. The solution to this scenario is to use vibro-audio interface where, *"Imagine the same scenario where the student has a touch-based device with vibro-audio interface and can trace the graphic element just by using haptic cues without getting distracted from listening to lecture"*. This issue led to the following research questions,

- *Is it possible to gain relation and quantity information using vibro-audio interface?*
- *Is it possible to perceive curves and angular lines using vibro-audio interface?*
- *Is it possible to get shape information through vibro-audio interface?*

Three conditions were designed to address these three questions respectively. The first condition assessed the relative relations and global structure of a bar graph, the second condition evaluated pattern recognition via a letter identification task, and the third condition investigated orientation discrimination through a geometric shape matching task. Though the scenario was based on a geometry class, these three questions and their results are also applicable for maps. Because, the map components such as corridors, walk ways are similar to bar graph, the junctions and intersection are familiar patterns as English letters and each map object will have an orientation with reference to its frame similar to that of geometric shapes. However, all the graphic elements used in this study are designed with an assumption that it will be smaller than the screen size of the display device, but maps are generally bigger and with the minimum line width being 0.35 inch, they cannot be fitted within the screen of a touch-based device as they are limited in size. This brings up the question of whether the interface will perform in the same efficient way if the graphics are bigger than the screen size. The only way to read a graphic bigger than the screen size is by panning. In visual interfaces this function can be performed easily as the visual component creates references in the image and allows the user to track the movement of the graphic material. But from pilot studies, it was found that this cannot be performed in the same way for non-visual interfaces as there is no clear reference point of the graphic that the user can follow and thus results in user being lost or disoriented after performing the function. This issue led to the following research question,

- What is the most efficient method to learn a map with non-visual panning?
- What kind of gestures can be used for non-visual panning in map learning?
- Is it possible to gain global spatial knowledge of the map by panning?

A human behavioral study was performed to answer these questions. The study was designed to compare map learning using various panning techniques. Even though the maps will be bigger than the screen size, it is possible that not all the map element would be bigger than the minimum width of 0.35 inch. Landmarks are critical for navigation as they form a solid reference point for the developing mental representation of a space. In most indoor settings, these landmarks will be of smaller size than the minimum width of 0.35 inch. For instance, imagine an indoor floor layout where the corridor is of 0.35 inch, then it is obvious that a water fountain will be smaller than that. So, it becomes untraceable for someone who is searching for water fountains. In such scenarios it is necessary to make it bigger and after perceiving the landmark, the map has to be reduced to its original size else other objects will not be perceivable as they will be much bigger in size. In order to make the map objects bigger or smaller, zooming functionality has to be implemented. In a visual setting zooming can be performed in multiple ways such as pinching, scrolling and using buttons. However, from pilot studies we found that these techniques cannot be implemented into non-visual map interface as there is no prominent reference point for the user that can convey the amount of zoom action performed. This challenge led to the following questions to be addressed experimentally:

- What is the most efficient way to learn a map and its objects with non-visual zooming?
- What kind of gestures can be used for non-visual zooming in learning a map object?
- What level of information should be presented at each zoom level?
- Is it possible to gain knowledge of landmarks of the map by zooming?

To address these questions, a human behavioral study was conducted similar to the panning study. A navigational game was designed where the goal of the game was to gather treasure from different treasure points by zooming in and zooming out.

The three studies focused on whether the interface assists in developing accurate mental representation of maps and their elements. Though the studies tested the efficacy in development of cognitive map and its subsequent map based tasks, it was all in a controlled setting and they did not address whether the information can be transformed to real-world tasks. Also the tasks did not involve any physical navigation. However, the ultimate reason for a visually-impaired person to use a map is to perform some navigational task such as way finding or free exploration. This brings up the questions of whether the interface is equally efficient in developing cognitive map such that it can be mentally transformed to assist real-world navigational tasks. This led to the following research question,

- Is it possible to transfer the virtual mental map learned offline by the interface for real-world tasks?
- How efficient is the interface in assisting real-world navigational tasks?

Thesis Statement

The goal of this work is to prove that it is possible to access and learn graphical material such as maps and graphs using the newly developed vibro-audio interface. The primary hypotheses are,

- For an unknown indoor environment, a Vibro-Audio Interface assists non-visual navigation better than the traditional hardcopy interfaces.
- Accessing larger and dynamic maps using the vibro-audio interface is efficient and easy than traditional hardcopy maps.
- Spatial knowledge of an indoor layout obtained through Vibro-Audio Interface leads to accurate spatial representation in memory and will result in better mental transformations when compared with the learning of same indoor layout using traditional hardcopy displays.

Experiment 1

In experiment 1, three conditions were conducted that assess comprehension of the relative relations and global structure between elements on a bar graph, pattern recognition via a letter identification task, and orientation recognition of complex geometric shapes on a shape discrimination task. Each experiment represents a different set of behaviors that rely on accessing an accurate spatial representation built up from learning common graphic material. They all compare two display mode conditions, one that employs the vibro-audio tablet interface at learning and another that employs hardcopy tactile stimuli produced by a graphics embosser (the gold standard for tactile output).

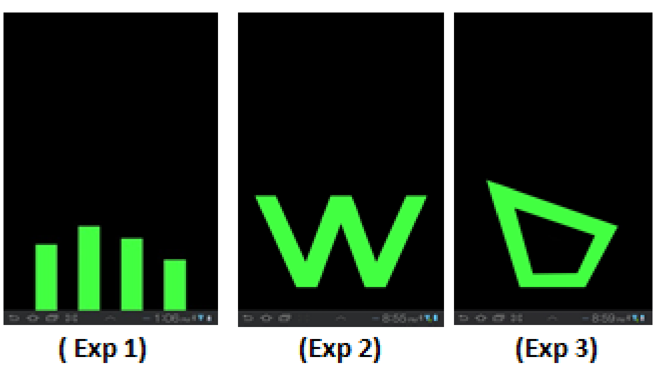

Figure 1. Example stimuli displayed on the touch-based device with the vibro-audio mode for the three experiments. Analog hardcopy tactile stimuli (not depicted) were used as a comparison in each experiment.

Figure 2. Subject tracing stimuli displayed on the touch-based device with the vibro-audio mode.

Method:

Participant Sample: The participants were recruited selectively as the system is targeted for blind and low vision users. The sampling involved three pre-requisites namely, participants should not have any known sensitivity problems in finger digits as the system's primary cue is Haptic, participants should not have any hearing impairments as the system partly works on audio cues, participants should not have any motor disability as the system involves kinesthetic feedback. The study was conducted with 12 sighted (six males and six females, ages 18-35) and 3 blind participants (2 males and 1 female, ages 22-38). All the three blind participants were congenitally blind and had no more than light perception. Even though the target is on blind and low vision users, the inclusion of blindfolded-sighted participants is considered as a reasonable sample as we are testing the ability to learn and represent non-visual material which is equally accessible to both groups.

A trial in each condition had two phases namely a learning phase and a testing phase. In learning phase the subject will learn the given stimuli in either of the display modes and in testing phase subject will answer questions based on the perceived stimuli. In bar graph study, a re-creation task was conducted in addition to answering task where subjects were asked to re-create the stimuli perceived in learning phase. A practice phase was conducted before each condition where the experimenter explained the task, goal and strategies of the conditions. The participants were then allowed to explore the practice stimuli and the experimenter provided corrective feedback at the end.

Independent Variables: Two display modes (Vibro-Audio display and Braille display).

Dependent Variables:

1. Performance accuracy in each task. (e.g., accuracy in identifying correct letter, shape, orientation accuracy in recreating the perceived object)

2. Learning time (Time taken to learn particular stimuli in learning phase)

Experimental design: A within subjects design was used in the experiment, with participants learning and testing in each of the two display mode conditions. The trials were randomized within each display mode and the order of display mode was counter balanced between subjects.

Preliminary Results:

Results from the experiment showed that a vibro-audio interface can lead to accurate learning of the graphical material. The data shows similarity in performance between the vibro-audio interface and traditional hardcopy display suggesting that both are functionally equivalent. The learning time was much higher for the vibro-audio interface when compared to hardcopy display, as it is easier to track the embossed braille stimuli. Also, the letter recognition performance of sighted blindfolded participants was significantly worse with vibro-audio interface (88%) when compared to the hardcopy display (100%). This is due to the impoverished orientation cues in the smooth touch-based display which made it harder to detect line orientation. But in hardcopy display, the embossed line makes it easier to detect orientations. However, blind participant were able to recognize all the letters in both display modes. For both 2nd and 3rd conditions, subjects self-reported difficulty in detecting oriented and curved lines as the cues were not sharp in smooth touch-based display when compared to embossed braille stimuli. In all, the vibro-audio interface proved to be functionally equivalent and better in some cases than the traditional hardcopy displays.