

Enhancing Independence for People With Low Vision to Use Daily Panel-Interface Machines

Xi Wang, Xi Zhao, Omprakash Gnawali, Weidong Shi

Department of Computer Science, University of Houston, Houston, TX 77004, U.S.A

Email: {xiwang, xizhao1, gnawali, larryshi}@cs.uh.edu

Abstract—Those that suffer from macular or vision degenerative diseases report lower levels of ease and ability to participate in daily activities, even when using the most advanced corrective lens technologies. They desire to gain independence in daily tasks, including using everyday machines. Plenty of these machines, such as microwaves, vending machines, and elevators, have a control panel as a user interface. Some have Braille (i.e. bumpy dots) on the panel for people with low vision to recognize. However, many of them have been designed without considering low-vision people’s convenience. Though many techniques have been proposed to improve the usability of these panels for people with low vision, they are difficult to deploy due to the high cost and low financial benefit. In this paper, we propose a panel reading assistive system which is wearable instead of being deployed on the machine’s side. The conceptual system instructs the user to operate on the panel of an everyday machine. It creates a plan according to the user’s need, and guides the user to implement it step by step. The user moves the finger under the system’s instruction via audio channel. Finally, the experiments in real life scenarios have demonstrated the feasibility and effectiveness of the system.

Keywords—*Low vision, Wearable panel reading system*

I. INTRODUCTION

Low vision is defined as chronic disabling visual impairments that cannot be corrected with glasses, contact lenses, or medical or surgical treatment. Low vision can result from a variety of diseases, disorders, and injuries of the eyes. People with vision acuity under 20/70 (standing 20 feet from an object to see it with the same degree of clarity as a normally sighted person could from 70 feet) are considered to have low vision. An estimated 3.5 million Americans have low vision; out of that group, approximately two million have mild low vision, which mainly affects driving and reading ability, and about one million meet the legal criteria for blindness. Across all age groups the incidence of more severe forms of low vision (including blindness) is about 250,000 new cases per year [1].

Everyday tasks are challenging for people with low vision, especially those with severe low vision or blindness. In early 2011, Action for Blind People carried out a questionnaire survey. The aim was to find out what issues are important to these people and what services and support they feel that they need or would like to receive [2]. One of the top five wishes identified was ‘Need to become more independent’. Unfortunately, plenty of private or public environments are not comfortable or functional for individuals who are blind or visually impaired. They cannot benefit all users of a facility. Therefore, researchers have taken much effort to develop

assistive technologies which can promote independence during daily activities for people with vision impairment. It is quite important for them to interact with home appliances (e.g., microwaves, washing machines) and public electric machines (e.g., vending machines and elevators) to live an independent life. However, to use them, people need to interact with their panels.

A panel is a flat, often vertical, area where control or monitoring instruments are displayed. They are found in factories to monitor and control machines or production lines, and in places such as nuclear power plants, ships, aircraft and mainframe computers. They are also widely used in private or public environments to manipulate household appliances, electronic devices or facilities, such as television, computers, and elevators. Panels are most often equipped with push buttons and analog instruments, whereas nowadays in many cases touchscreens with flat, smooth surfaces are used for monitoring and control purposes. The panel mentioned in the paper refers to a user interface of a daily machine. Panels sometimes have Braille next to the panel button. However, as seen in Fig.1, lots of panel-interface machines don’t offer any special assistance (i.e., tactile feedback) to people with low vision. As a result, they either have trouble using these electronic devices or public facilities, or cannot use them at all.

In order to reduce the usage barrier, there are a few solutions for people with low vision. Three prevailing solutions are talking [3][4][5], tactile feedback [6][7][8], and gesture based [9][10][11] interfaces. For instance, a popular way for them to interact with home appliances is to ask others to attach a few different salient markers on each button of the interface as landmarks. However, the marker solution can only be used in private places. Other solutions need to either embed new functionality into the current machines which are originally designed for fully sighted people, or develop a completely new panel-interface for a machine from scratch [9]. They have demonstrated effectiveness. However, most of them have one crucial drawback, which is because new design or improvement on the machine side is required, people need to replace the current machines. The cost for update or replacement prevents the spreading of these solutions.

Unlike current solutions which are on the machine side, our system proposed in this paper is on the user side. It requires no update or replacement on the machine side, and is portable; thus it can be useful to more than one kind of machine. It’s a wearable panel reading system which can provide a speech interface to advise the hand actions for smooth use of the



Fig. 1. Left two are panels of elevators and right two are panels of vending machines. Top two panels have Braille, that is circled by red loops. The Braille provides tactile feedback to blind people. However, many machines have no special interface for them, like those in the bottom two figures). Blind people have met with great difficulties during the use of these panels.

machines. The system perceives the panel using a camera, interprets the steps with the user's intention, and provides speech command for the user. Aided by the system, people with low vision can interact with the panels even though the machines provide no special design for them. The user first tells the system what she/he needs from the machine with a short pre-defined command. Then the system creates a plan based on the command. Using this plan, the system step-wisely advises the users' hand action. A pure color cap is attached or nail polish is painted on his/her fingertip. The system can thus find the user's fingertip by detecting the color marker with a color detection algorithm. The user moves her/his fingertip on the surface of the panel under audio direction from the earphone, until her/his fingertip is on the top of the right button. There are two essential contributions of our system: first, our system is wearable and unobtrusive, making it possible for the users to bring it during their daily activities. Second, the users do not have to passively wait for the low-vision-people-unfriendly panels to be redesigned.

II. SYSTEM DESIGN

A. Overview

As seen in Fig.2, the system consists of two major components: the client and the server. Each has its computation module, which consists of any component that is powerful enough to fulfill the computation need. Other than the computation module, the server side includes the database (information of all panels). It saves storage space and complements inadequate processing power on the client side. The client side includes the user interface module, the signal capture module, where the user is directly involved. The signal capture module includes two sensors that collect user data for measuring and recognizing the user's behavior. The user wears a mounted camera, which is connected to a smartphone. This camera can be tiny and unobtrusive, like a micro camera as small as a coin. The purpose of using the small camera is to avoid extra

burden on the user in terms of weight and social attention. The camera captures visual content and sends it to the computation module of the client side (i.e. a smartphone). The user interface module includes an earphone and a smartphone. The user starts or stops the service of our system through the smartphone. The user gives the audio commands such as 'buy a bottle of Pepsi', so the system can know what the user intends to do, and create a plan to guide the user. The smartphone then gives the user instructions via audio feedback, which could come via earphone or embedded speaker on the smartphone. On the server side, its computation module decides which machine the user would like to use, and sends a serial of data regarding the machine to the client side. Based on the data and the position of the user's fingertip, the smartphone can create instructions for the user.

We have utilized several technologies to implement our design, such as an indoor positioning system (IPS), Scale Invariant Feature Transform (SIFT) [12], color detection algorithm [13], speech recognition [14] [15], and planning algorithm [16]. IPS is a network of devices which are able to wirelessly locate objects or people inside a building. We are familiar with IPS's peer Global Positioning System (GPS), which provides information about location and time. The difference between them from the user's perspective is that IPS can track you to a specific spot inside a building, while the indoor position given by GPS is quite inaccurate. We use IPS to track the user in order to decide which machine the user will use. SIFT is a method to detect distinctive, invariant image feature points, which can be easily matched between images to perform tasks such as object detection and recognition, or to compute geometrical transformations between images. SIFT is rotation-invariant, which means that even if the panel is rotated and transformed across different frames in videos, we can find the same image feature points and thus align the panels in different frames into a normalized one, such as the frontal panel. We use it to recognize if the area that the user is touching is the button that the user is supposed to press. Color detection algorithm can find the areas in a image which contain the designated color within a certain range. We use it to detect the user's fingertip. Speech recognition is to translate speech into text, so the computer knows the meaning of the user's command. The planning algorithm is to create a step-wise plan based on the user's command. We utilize the planning algorithm we have proposed in [16]. It navigates the user through complex tasks and includes a novel canonical Priority Graph (PG) to model the complex activities and implement the planning function.

B. Hardware

As seen in Fig.3, the hardware includes an earphone, a camera, and a smartphone. The camera is Looxcie LX2 Wearable Video Cam for iPhone and Android. It's light, hand-free, and small enough to attach anywhere on the user. Its built-in WiFi can send images or video to the smartphone. It has an adjustable design for optimizing fit and comfort. The user can use any earphone that she/he wishes. Nowadays, tiny, light earphone is common in daily life. As long as it speaks instructions, the user can choose anything. In cases that the user forget to connect the earphone, the built-in speaker in the smartphone can also be used as the speech interface.

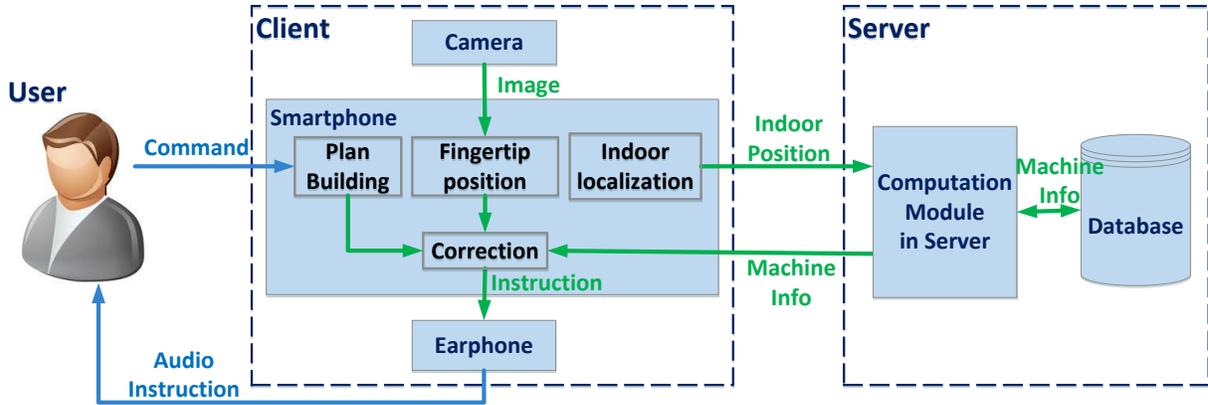


Fig. 2. The flowchart of the system. The server side contains the database and its computation module. It sends the information of the machine to the client side. It includes the panel image, the button positions on the panel, and the semantic meaning of each button. The client side listens to the user's command and builds a plan, following which the user can make the machine work as per the command. The camera detects the user's fingertip and calculates its position on the panel. The smartphone compares the area the user is touching on and the original image of the panel, and then calculates the direction the user's fingertip should move towards.

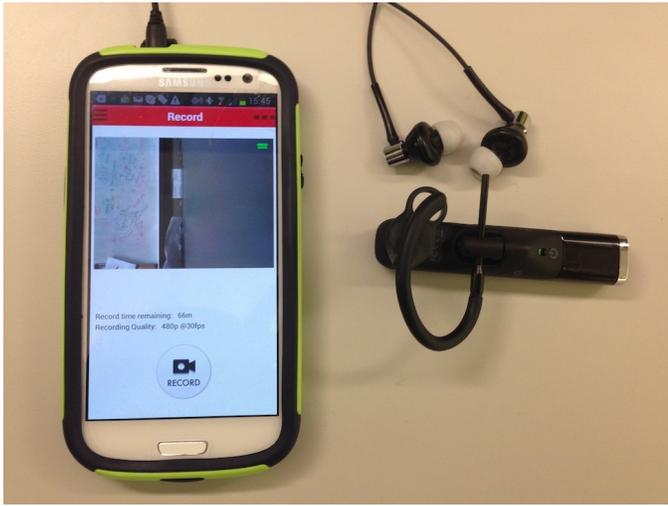


Fig. 3. Hardware of our panel reading system. It includes an earphone, a camera (on the right), and a smartphone. The camera captures the visual content in the scene and transmits the videos via the wireless connection. The visual content can be displayed on the smartphone's screen in real-time. The smartphone detects user's position, analyzes the user's action by processing the image, decides the user's next move, and gives user instruction through the earphone.

Smartphones possess technological advances in computing and I/O capabilities as well as network connectivity [17]. All the sensing and computing technologies of this solution are available in common smartphones. Normally, a smartphone contains several useful sensors (e.g., touchscreen, accelerator, gyroscope, microphone, magnetometer, GPS, camera). The indoor positioning system we can use is based on the dead reckoning positioning method, 2D barcodes, and data from smartphone's sensors (accelerators and magnetometers) [18]. For user's convenience, we decide not to use the built-in camera because it requires the user to adjust the entire phone instead of a tiny camera. The smartphone can not only localize the user, but also conduct image or audio processing (SIFT,

color detection, speech recognition).

C. Software

Server: On the server side there is a database, which manages the information for each panel. The system retrieves the information based on the indoor position given by the client side. The panel information includes panel image, button positions on the panel, and meanings of each button. This database is also collected and built based on the IPS network. On our campus, for example, we can build the IPS network first and label the position and the type of each vending machine within the IPS network. Note that the IPS network can cover other machines in the public space not limited to vending machines. Specifically, to construct this dataset, the panel position and items that each machine supplies are labeled; pictures of each panel are taken, where the panel areas are cropped and saved. The semantic meaning of each button is interpreted mainly in consistency with the requirement of the planning algorithm. Taking one machine for example, it has A-C, 1-9, and start buttons on its panel. According to the instruction sticker on the machine, after the payment, the user should press a letter, then a number, and last the start button. These steps are learnt by the planning algorithm for the online use. When the client side requests a panel information with an indoor position, the server would send it to the client.

The computation module in the server determines which machine the user intends to use by calculating distance between the user's indoor position and each nearby machine's position. The machine with the smallest distance is chosen. Then the server sends the information about this machine to the smartphone.

Client: The user carries devices of the client side: smartphone, camera, and earphone if necessary. When the user finds a vending machine, she/he starts the system. The system can localize user's position using IPS. After the client system received the information from the server, it tells what this



Fig. 4. Real settings: on the left is the panel of the vending machine in a classroom building, on right is the panel of the microwave in a public lounge on campus. The buttons on the vending machine panel bulge, while those on the microwave are basically flat though the protective film is worn out a little.

machine can supply (e.g., soda, water, snacks). If the user is interested, she/he communicates with the system via oral command on the item that s/he wants. Then the system sends this information to the server for the plan and then step-wisely advises the user.

The user can get the item if following the plan. The first step is to find the panel. The system instructs the user the coarse position of the panel on the machine, e.g. ‘on right’, ‘on upper left’. The appearance of the panel is also recorded in the database. Normally, the panels on vending machines are not a smooth surface. It’s easy for users to find given the coarse panel position on the machine. Some machines have flat panels like those of microwaves. However, given the coarse position, the system can easily direct the user to the panel, since the interface side of a microwave is much smaller.

The plan includes pressing a certain sequence of buttons, like A-F or Start. The user’s fingertip is covered with marker (colored loop or nail polish), which is a unique and constant color to prevent the color from being similar to the panel color and not allowing the fingertip to be recognized correctly. When the user holds the position of her/his fingertip on the panel, the system detects the position of the fingertip. First, the smartphone finds the panel in frames using SIFT. It then finds the fingertip by detecting the color of the specific marker on the panel area. When the smartphone finds out the position of the user’s fingertip, it calculates the vector from the center of the fingertip to the center of the target button. The vector is translated into directional audio instruction. When the finger is on the right button, the system asks the user to press it.

III. EVALUATION

We designed a proof-of-concept study to test the effectiveness of the panel reading system in real settings as seen in Fig.4. We manually loaded the information of these two machines’ information to the database on the server side.

We recruited five participants without any vision impairment. Each of them wore a professional low vision simulator [19], a goggle-like product which can generate moderate

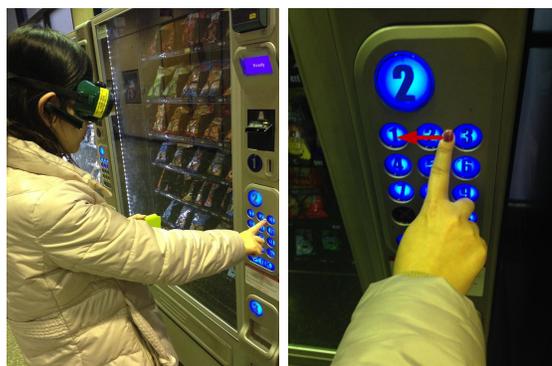


Fig. 5. The participant operates the panel with the help of our system.

low vision (20/160 acuity) for normal people. The participants used both the microwave and vending machine. If a participant can not successfully complete the task on the machine in three minutes, her/his completion score is set to zero. Otherwise, the completion score is set to one. They were operating the two machines wearing the devices of the client side in the first round. Once our system was set up and initialized, we started measuring the performance. The second round, they did not use our system. A participant had two different tasks to accomplish in two rounds on one machine in case that the participant remembered the position of certain buttons and the sequence of steps.

We know that even if a participant randomly presses buttons, the microwave might start or the vending machine might pop out an item. So if the machine does not work as the participant pre-claims, then we still see it as non-completion. For example, a participant says she/he would like for the microwave to warm for 30 seconds, but the food turns out to be warmed 3 minutes, or she/he wants some chips while the vending machine outputs a bag of cookies.

Figure 5 demonstrated how our system helps the participant use the vending machine. In the left figure, the participant’s eyes were occluded by the goggles to simulate low vision. On the participant’s left hand was the smartphone. The earphone and the tiny camera were respectively worn on the left and right ears of the participant. The camera captured images from the first-angle view of the participant which was pointing to the vending machine. This vending machine required the participant to swipe a credit card first and then press the number of the item she/he wanted. The slot of the credit card scanner was easy to find and use. In the right figure, the participant was trying to find the target button. She was supposed to press button 1. The system detected her fingertip by finding the color of her nail polish on her forefinger. It calculated that she should move her finger to left. After she completed all the number combinations which correspond to the item that interested her and pressed the OK button, she could pick it up.

In Fig.6, we can observe that under instructions of the panel reading machine, all five participants successfully got what they wanted, warmed their food up in the microwave, and purchased items supplied by the vending machine. However,

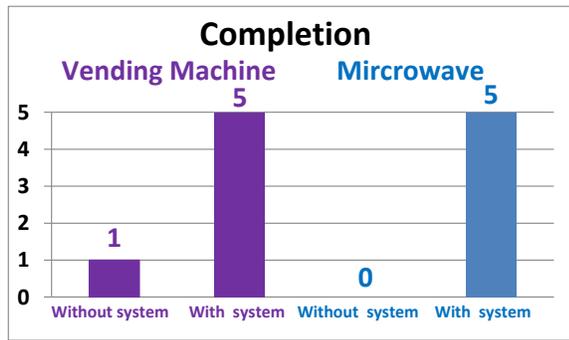


Fig. 6. With the panel reading system, all five participants were able to get both vending machine and microwave to work. Without it, most of the participants could not find the right buttons.

without the aid of the proposed system, all of them failed in using the microwave and four out of five failed in using the vending machine correctly.

IV. CONCLUSION AND FUTURE WORK

We propose a panel reading system targeting on helping people with low vision to use panel-interface machines with a real-time interactive guidance. Unlike fully sighted people, the users are not able to read the button on the panel without additional assistance (i.e., Braille near the buttons). Sometimes, if the panel is flat and smooth, they cannot even easily find where the panel is. The system functions as an augmentation of the user's ocular cognition. It guides the user to find the panel and to operate on the panel step by step. Our experiments demonstrate the effectiveness of the system to build independence in terms of use of the common machines.

We view the work described in this paper as only the beginning. We will implement IPS and test its performance of localizing the machine which the user would like to use; build the system out to be able to discover new types of panels (e.g., panels on washing machines or ATM), and provide instructions based on the type (e.g., all vending machines have similar order of operations); and recruit people with low vision in order to evaluate user perception of the system (i.e., is it actually building independence for the user).

REFERENCES

- [1] R. W. Massof, *Low Vision and Blindness: Changing Perspective and Increasing Success*. National Federation of the Blind, 2006.
- [2] S. Wright, C. Vince, and S. Keil, *Tell us what you want*. Action for Blind people, 2011.
- [3] C. S. Ltd, "Talking combination oven," 2012, https://secure.vvhosting.co.uk/cobolt/store/eric_viewItem.asp?idProduct=248.
- [4] "Maxi-aid," 2013, <http://www.maxiaids.com/store/default.asp>.
- [5] "Ivee digit," 2013, <http://www.assistech.com/cgi-bin/shopper.cgi?preadd=action&key=703538>.
- [6] SSMR, *RNIB Understanding the Needs of Blind and Partially Sighted People: their experiences, perspectives, and expectations*. University of Surrey RNIB, 2009.
- [7] *A kitchen for you*. Disabled Living Foundation, 2002.
- [8] M. Scheiman, M. Scheiman, and S. G. Whittaker, *Low Vision Rehabilitation A Practical Guide for Occupational Therapists*. SLACK, Incorporated, 2006.
- [9] F. E. Sandnes, T. B. Tan, A. Johansen, E. Sulic, E. Vesterhus, and E. R. Iversen, "Making touch-based kiosks accessible to blind users through simple gestures," *Universal Access in the Information Society*, vol. 11, no. 4, pp. 421–431, November 2012.
- [10] S. K. Kane, J. P. Bigham, and J. O. Wobbrock, "Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques," *ACM SIGACCESS Conference on Assistive Technologies*, pp. 73–80, October 2008.
- [11] D. Mcgookin, S. Brewster, and W. Jiang, "Investigating touchscreen accessibility for people with visual impairments," in *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*, October 2008, pp. 298–307.
- [12] R. Hess, "An open-source siflibrary," in *Proceedings of the international conference on Multimedia*, New York, NY, USA, 2010, pp. 1493–1496.
- [13] R. P. Schumeyer and K. E. Barner, "A color-based classifier for region identification in video," in *Visual Communications and Image Processing*, San Jose, CA, USA, 1998, pp. 189–200.
- [14] L. Rabiner and B.-H. Juang, *Fundamentals of speech recognition*. NJ, USA: Prentice-Hall, Inc., 1993.
- [15] D. Jurafsky and J. H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. NJ, USA: Pearson Prentice Hall, 2000.
- [16] "Anas: An activity navigation assistive system using a priority graph and multi-modality sensing," *In ACM Transactions on Interactive Intelligent Systems (Under revision)*, pp. 1–21, 2013.
- [17] L. Wood, "Worldwide smartphone markets: 2011 to 2015," Research and Markets, USA, Tech. Rep., May 2011.
- [18] A. Serra, D. Carboni, and V. Marotto, "Indoor pedestrian navigation system using a modern smartphone," in *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*, New York, NY, USA, 2010, pp. 397–398.
- [19] "Low vision simulator," <http://www.lowvisionsimulators.com/>.