Portable Form Filling Assistant for the Visually Impaired

En Peng

e.peng@curtin.edu.au

Patrick Peursum

Department of Computing, Curtin University of Technology
GPO Box U1987, Perth, Western Australia, 6845
p.peursum@curtin.edu.au

Ling Li

l.li@curtin.edu.au

Svetha Venkatesh

s.venkatesh@curtin.edu.au

The filling of printed forms has always been an issue for the visually impaired. Though optical character recognition technology has helped many blind people to 'read' the world, there is not a single device that allows them to fill out a paper-based form without a human assistant. The task of filling forms is however an essential part of their daily lives, for example, for access to social security or benefits. This paper describes a solution that allows a blind person to complete paper-based forms, pervasively and independently, using only off-the-shelf equipment including a Smartphone, a clipboard with sliding ruler, and a ballpoint pen. A dynamic color fiducial (point of reference) marker is designed so that it can be moved by the user to any part of the form such that all regions can be "visited". This dynamic color fiducial marker is robust to camera focus and partial occlusion, allowing flexibility in handling the Smartphone with embedded camera. Feedback is given to the blind user via both voice and tone to facilitate efficient guidance in filling out the form. Experimental results have shown that this prototype can help visually impaired people to fill out a form independently.

Visual marker design, visual marker recognition, mobile devices, mobile HCI, fiducial recognition, user studies

1. INTRODUCTION

Every now and then we are asked to fill out forms. Despite the increasing availability of forms in electronic format, we still need to fill out hardcopy forms in many circumstances, for example, due to the requirement of handwritten signatures. While blind people are able to 'read' most hardcopy forms thanks to the development of optical character recognition technology, it is still very difficult for them to locate the form fields printed on the plain paper. In order to complete a hardcopy form, assistance from a sighted person is the only option for those with total vision loss. However, after consultations with blind experts from the Association for the Blind of Western Australia, we have found that the involvement of a third person introduces important social issues - feelings of anxiety while searching/waiting for someone to help, feelings that their privacy is being invaded, and concerns over whether the assistant can be trusted [15].

Figure 1: Form filling examples: (a) original images (320 x 240 pixels) observed from the embedded camera on Nokia E71 Smartphone; (b) illustration of computed visible form regions; (c) enlargement of the form near the pen tip position.

To enable the blind people to fill out a form independently, the form must be accessible. The standard accessible text representation widely used by the blind is Braille, which encodes the text into a binary code and represents them as raised dots through an expensive Braille embosser [11]. Although forms represented in Braille format may be available occasionally, they do not provide information regarding the layout of the form's fields. Most importantly, they are usually for reference

only and cannot be submitted, e.g. accessible tax forms which are provided by some government agencies[5]. In order to make form fields accessible, some researchers [4] have used a Braille embosser together with a fusing device, e.g. the Picture in a Flash (PIAF) Tactile Graphic Maker [11], which can swell the black printed regions on a specially-made swelling paper. This would require a reprint and modification of the original form which might not be feasible in most circumstances, and thus faces all the same barriers to wider availability as Braille. To our knowledge, there are no existing devices that enable blind people to fill in a form printed on a piece of ordinary paper.

To avoid reprinting the form and to allow the visually impaired to understand/write on the form, computer vision based techniques can be used to analyze the form. Under certain constraints, existing computer vision technologies can be adapted to recognize the form information (text and layout) and detect the pen tip position quite reliably. The remaining problem is to recover the form pose so that the pen tip position on the form can be computed in real time. To perform pose recovery efficiently, there are two possible ways: 1) using the natural features of the form and 2) using pre-defined fiduciary markers. If there are sufficient distinctive features [16] on the form, it is possible to recover its pose with prior training. However, this can take a few hours making it impractical for impromptu use. In addition, the trained features may become obscure as form fields are filled in, leading to estimation failures. Therefore, fiduciary markers are usually adopted [9, 12] as they can provide a sufficient number (at least four) of reliable point correspondences between object coordinates (in our case, the form) and captured image coordinates (See Figure 1). As the layout of the form is unknown, the fiduciary marker(s) cannot be fixed at pre-defined location(s) within the form as it may block a form field. However, if the fiduciary marker is positioned outside the form, the camera must be far enough to observe the entire form and the markers, restricting possible camera poses. Moreover, since we wish to use a mobile device, low image resolution is required for real-time video capture. This further makes it impractical to handle far field views while maintaining accuracy. Therefore, the main challenge here is to make the pose of the mobile camera more flexible, whilst recovering the pose and retaining the ability to write on any region in the form.

This paper proposes a solution that enables blind people to fill a hardcopy printed form, using a standard pen in a pervasive manner, using an ultra-portable Smartphone and an ultra low cost off-the-shelf clipboard set. A dynamic fiduciary marker is designed to robustly calculate the camera pose: the camera can be placed to observe the complete form

or only a small region inside the form depending on where the user moves the marker (see Figure 1 for examples). A regular pen is used by the user for guidance. As the target user of this tool is expected to be totally blind, feedback methods have been specially designed to ensure that a sufficient portion of the fiduciary marker is observed for pose estimation and the user, via their pen, is guided to the target form field. The user may cycle through target fields. The prototype of the proposed system has been evaluated with four totally-blind volunteers who were asked to fill in a test form. Results are based on the accuracy of the volunteer's ability to fill in the test form as well as a survey of their opinion regarding the system. Favorable feedback has been received for current interactions from those participants.

The main contributions of this paper include:

- The design of a form filling assistant prototype for people with total vision loss using low cost off-the-shelf products.
- The design of a dynamic fiduciary marker, which allows arbitrary positioning/writing on any part of the form, which can be effectively extended to a sheet of paper with any content, enabling flexibility in camera pose.

The remainder of this paper is structured as follows: Section 2 gives the background on related work; Section 3 explains the positioning on form; Section 4 describes the prototype setup and usage; while Section 5 shows the experimental results before Section 6 concludes this paper.

2. RELATED WORK

In this section, related work which are based on forms on an ordinary sheet of paper is introduced.

2.1 Optical Character Recognition

Optical Character Recognition (OCR) technology, which extracts text from images, enables visually impaired people to obtain printed, typewritten or even handwritten text. The obtained text can be processed into accessible media such as voice via text-to-speech(TTS) engines or Braille via a Braille embosser [11].

It has been almost a century since the first OCR system was invented for statistic purpose [3]. Up until now, there are a wide range of OCR software for different languages, such as ABBYY FineReader, Kurzweil, and Adobe Acrobat. Such software normally produces highly accurate results for typewritten/printed text, but the recognition of handwritten text is still an active research topic. The common usage of OCR technology involves three

steps: first, the document is digitized using a flatbed scanner attached to a computer; second, the page layout is analyzed before the text is recognized using OCR software; third, the text is spoken aloud via a TTS engine. A combination of digital camera with laptop computer, or a camera embedded in a Smartphone has made OCR technology far more portable.

Although OCR technology allows a blind person to 'read' the text in the scene, it does not allow them to locate the text in the scene. Therefore, OCR technology is not sufficient for helping the blind to fill in a printed form. To achieve that, assistance in positioning on the paper-based form is required.

2.2 Positioning on a Planar Surface

Many physical display media containing text information are planar, e.g. paper, whiteboard, and screen. Unlike Braille where raised dots are used, these media do not enable blind people to 'feel' the information by touch. There are two main positioning devices for planar surfaces: electronic and vision based.

2.2.1 Electronic Positioning

The most common electronic based device is the mouse attached to a computer, which allows accurate positioning on any pixel on the screen. This position can be provided as feedback to the blind user via speech, e.g. speaking the 'hovered' text. Alternatives to a mouse such as light pen, touchscreen, or hybrid interactive surface [7] allow the user to point object(s) or finger(s) on the screen directly to locate desired position(s). But a blind user would not know its physical position on the screen, hence such a system only works in digital form. The electronic pen is an alternative device which enables the user to hold a pen-like object working on a tablet, which senses the pressure of the digital pen. A digital notepad with a digital ink pen [14] already exists on the market which allows the recording of pen on the tablet in electronic format while the ink from the same pen is printed on the paper on the tablet. One limitation of these electronic based devices is that they all require expensive custom-built hardware, and none has attempted to assist blind users.

2.2.2 Vision-based Positioning

Vision based pointing devices are generally more affordable, as the consumer market is overwhelmed by low cost cameras and products with embedded cameras. To enable positioning using a camera, the pose must be estimated before the pen position is detected.

Fiducials, a.k.a. fiduciary markers, are usually placed on the plane for pose estimation [6, 9, 10, 12, 13, 17]. Common fiducials include point fiducials and planar fiducials. The point fiducial [13] represents a landmark point, while the planar fiducial provides multiple feature points [10]. The planar fiducial can be either hard coded or customized with natural features [2] by a user. A planar fiducial or at least four point fiducials with known planar geometry (normally at four corners of a rectangular plane) are sufficient for estimation of the pose between a camera and a planar object. The 2D homography can be estimated relating an image point with its position on the reference plane. This 2D homography is then used to compute the position of the pen on the reference plane given its observation on the image. ALED (Light-Emitting Diode) pen is a reliable device for vision based positioning. The camera only needs to find the brightest point emitted from the LED on the pen tip. The position on the planar surface can be computed using 2D homography. Another solution can be found by using IR (Infra-Red) LED lights and a camera with an IR filter. The Wiimote Whiteboard [8], which utilizes the game controller of Nintendo's Wii, is a successful application. However, LED pens do not have visible ink.

3. POSITIONING ON FORM

In this section, we describe a computer vision approach for obtaining the pen tip position of a regular pen at any position on a form (or in fact any other thin planar media). A dynamic fiduciary marker is also designed to allow the usage of handheld moving camera with flexible pose/distance from the form. The computer vision algorithms used to detect the marker and computation of the pen tip position on the form are introduced.

3.1 Dynamic Fiduciary Marker

As introduced in Section 2.2, fiducials can be classified into point fiducials and planar fiducials. The point fiducials are usually very simple, e.g. circle. At least four points with known geometry are required for pose estimation. It is difficult to guarantee that these fiducials are always visible, especially when the viewfinder of the camera is not observed by the human. If many fiducials are put together, they must have identifiable features, e.g. color or pattern, which also makes recognition more complex. In addition, the fiducial also needs to accommodate the scaling problem, so that it can be detected no matter how far the camera is. Some fiducials are designed to have different scales, for example,

Fiducial	
Code	r, g, r, g, r, g, b, r, g, r, b, r, g, b, r, g, b, r, b, r, b, g, r, g, b, g, r, b, g, b, r, b, g, b, g, r, g, b, g, r, b, r, b, r, g, r, g
Position	01,02,03,04,05,06,07,08,09,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52

Figure 2: The linear fiducial used in our method

a multi-ring color fiducial with core circle and rings painted with different colors where the number of rings in the fiducial tells the fiducial level it belongs to [1]. However, as the number of levels increases to adapt to different scales, the size of the fiducial also increases and will become cumbersome.

We therefore design a dynamic fiduciary marker, consisting of three identical linear fiducials with the geometrical constraints that two of the parallel ones are perpendicular to the other, as shown in Figure 1(a).

Each linear fiducial consists of 52 linear uniformly spaced circular dots, filled with three basic colors (red, green and blue), as shown in Figure 2. Each neighboring group of seven dots defines a unique code that can be decoded to a unique position (with direction) on the linear fiducial. It is robust to occlusion: observing less than 14% of the linear fiducial is sufficient to confirm the existence of the linear fiducial and pinpoint position.

The dynamic fiduciary marker allows two parallel linear fiducials being moved in either direction along the other fiducial, as long as their geometrical relationship is retained. Therefore, the dynamic fiduciary marker does not block any part of the form permanently as the user can adjust the marker dynamically.

3.2 Form Pose Estimation

With the above dynamic fiduciary marker, it is possible to estimate the relative pose between the camera and the form. This involves three steps: detection of dynamic fiduciary marker, extraction of landmark points and pose recovery.

3.2.1 Detection of Dynamic Fiduciary Marker

To determine if the form pose can be estimated, we first need to detect if the dynamic fiduciary marker is in the image. As our dynamic fiduciary marker consists of three identical linear fiducials with known geometrical relationship, we need to extract automatically from the image the following data: the filled color circles used in the fiducial, up to three valid linear fiducials, and the geometrical relationship among the detected linear fiducials.

We do this by finding red, green and blue connected components and perform the Hough transform on the centers to find collinear dots. Up to three lines with valid color code (at least seven dots in length) satisfying projective invariance are obtained before their geometrical relationship is computed. In this way, the vertical coordinates for both parallel linear fiducials are estimated.

3.2.2 Extraction of Landmark Points

Since four points are sufficient to recover the relative pose of the camera and the plane, we choose two dots from each of the parallel linear fiducials. Since

the vertical coordinates of the dots are computed in the previous step and their horizontal coordinates are known by design, the coordinates of the landmark points in both world and image space are ready. With these information, the 2D homography between the image and the plane can be easily computed.

3.3 Pen Tip Localization

A pen tip may appear in different shapes in a perspective image. Training on its shape could be a solution, but it may not work well if a pen with different design is used. We therefore adopt a simple color based pen-tip detection method: assuming the pen has a red casing (the ink color may vary by using a refill with a different color), we only need to find the left most (right most for left-handed user) point of the largest red region projected on the form area (not on the ruler). We choose the red color for the casing of a pen because red color is least likely to appear on the form. More importantly, this allow us to speed up pen tip detection by taking advantage of the color information already processed during fiducial detection, an issue given the limited CPU resources of Smartphones.

4. PROTOTYPE DESCRIPTION

4.1 System Setup

The proposed system is composed of three off-the-shelf products: a typical mid-range Smartphone (Nokia E71) with a cradle, a clipboard with sliding ruler and a black ballpoint pen with a red casing. Whilst it is hard to find a black ballpoint pen with red casing on the market, we simply use a black ink refill for a off-the-shelf ballpoint pen with a red casing. In addition, we print out three identical linear fiducials and attach two of them horizontally on the ruler while the third vertically on the clipboard. A Velcro strip is also attached on the clipboard and ruler so that the ruler can be stabilized on demand. The objects used in this tool are illustrated in Figure 3.

Figure 3: Prototype Setup: (1) Smartphone; (2) Clipboard; (3) Sliding Ruler; (4) Black Ballpoint Pen with Red Casing; (5) Form; (6) Smartphone Cradle; (7) Velcro strip for stabilizing the ruler.

There are many ways for obtaining form data from a sheet of paper as described in Section 2.1, such as scanning/snapshot plus OCR software. Therefore, form data is assumed available and is not the focus of this paper. At this stage, all form data is assumed to be known a priori.

Table 1: List of Feedback Tones

Pitch	Low (300Hz)	Mid (500Hz)	High (700Hz)
Object	Smartphone	Pen	Ruler or Pen
Meaning	Pose Incorrect	Position Incorrect	Position Correct

4.2 Navigation on the Form

To allow the blind user to navigate the form, we require the user to memorize the meaning of three easy-to-distinguish tones with different pitches, as listed in Table 1. The low pitch tone indicates the user must adjust the Smartphone pose until the beeps stop; the mid pitch tone prompts the user to adjust the pen position until the beeps stop; and the high pitch tone is used to request the user to keep the ruler/pen in that position. After these three tones are memorized, the blind user may start a six-step procedure (transitions between steps are automatic).

Step 1. Selection of the target form field

Like GPS navigation systems, the target form field is first selected. As the information of the list of empty form fields is already available, left/right buttons can be used to select the form field to be filled. Once the button is pushed, the corresponding form field label is provided via voice feedback. In this way, the user can stop pushing buttons until the target form field label is read. No more buttons are required to be pushed in the following steps unless the target form field needs to be changed.

Step 2. Initial positioning of the Smartphone

Once the target form field is chosen, the user is requested to place the Smartphone to the left of the ruler via voice feedback ("Mobile phone. Move to the left of the ruler"). The goal of this step is to make sure sufficient portions of three linear fiducials are visible in the image so that the ruler position can be determined. Low pitch tones will keep beeping until the Smartphone is positioned at a valid position.

Once the beeps stop, the next step is followed to assist in positioning the ruler. See Figure 4(a) for illustration.

Step 3. Positioning of the ruler

With three visible linear fiducials, the vertical coordinate of the top edge of the ruler can be computed. The goal in this step is to guide the user to move the ruler so that the top edge of the ruler coincides with the bottom border of the target form field. Depending on the computed position of the ruler, instruction "Ruler. Move up" or "Ruler. Move down" is provided as voice feedback. The Smartphone may be required to move simultaneously with the ruler if low pitch tone beeps occur, as the position of the ruler cannot be computed in that case. Once the ruler is placed at the right position, the high pitch tone will start. The user then may adjust the ruler position until the high pitch tone beeps constantly. After keeping the ruler at the right position for half second, a voice feedback will be provided as "Ruler. Stabilize". In this case, the user can stick the Velcro strip on the ruler to the clipboard to stabilize the ruler. See Figure 4(b) for illustration.

Step 4. Positioning of the Smartphone

Once the ruler is stabilized, it is no longer necessary to observe all three linear fiducials since the vertical coordinates of the parallel linear fiducials are known and should not change. Therefore, in this step, the Smartphone needs to be placed at a position that observes the complete form field and the parallel fiducials. Since the form field is very close to the ruler, voice instruction "Mobile phone. Move along the ruler until beeps stop" is provided. When the low pitch tone beeps stop, the next step is to position the pen. See Figure 4(c) for illustration.

Step 5. Initial positioning of the pen

As both the Smartphone and the ruler are already in the correct positions, we just need to guide the pen to the beginning of the form field. The voice instruction "Pen. Move along the top of the ruler until beeps stop" is provided if the pen is not yet visible in the image. By moving the pen along on top of

Figure 4: Usage: (a) Step 2, moves the Smartphone to the left of the ruler (three linear fiducials can be at least partially observed); (b) Step 3, moves the ruler up or down (the top of the ruler can be aligned with the bottom of the form field) and stabilizes it using the Velcro on the left bottom of the ruler once it is in position; (c) Step 4, moves the Smartphone along the ruler (complete form field and partial parallel fiducials can be observed); (d) Step 5, moves the pen along the top of the ruler (until it reaches the beginning of the field).

the ruler, it will eventually become visible as the top edge of the ruler is at least partially visible. When the pen tip becomes visible, mid-pitch beeps will stop and corresponding voice instructions (“Pen. Move left” and “Pen.

Move right”) will be given. Similar to the positioning of the ruler, high pitch beeps will be given when the pen stops at the correct position. After beeping for half second, writing can start in the next step. See Figure 4(d) for illustration.

Step 6. Writing

To ask the user to begin writing, voice feedback “Pen. Start writing” is provided before the instruction is given as “You will hear beeps if outside the form field”. The blind user can start writing now and will hear mid-pitch beeps once the pen tip is outside any border (top, bottom, left or right) of the form field.

5. RESULTS

5.1 Participant

Four totally-blind people from the Association for the Blind of Western Australia volunteered to participate in the field trial in evaluating the proposed form filling assistant prototype. According to the order the trial, the participants are labeled as P1, P2, P3 and P4. It is notable that P2 is blind since birth and has not learnt writing in letters.

5.2 Task

Each participant is asked to complete a form with four blank fields as shown in Figure 5. Since the goal of this experiment is to test the positioning ability, four fields are not aligned either horizontally or vertically. The first form field is used for practicing during learning. The last three form fields are for test purposes. We only told the participants that there were four form fields on the form and they could write anything in the form field. No further information was given to the participant regarding the position of each form field.

5.3 Procedure

On arrival of each participant, the objects utilized in this tool (the Smartphone with cradle, the clipboard with ruler and the pen) are first introduced to the participant by touch. After the participant is able to

identify different objects, the participant is asked to complete the first form field following the voice and tone feedback from the Smartphone as introduced in Section 4.2 under supervision. It can be repeated until the participant feels comfortable about the usage. Next, the participant is asked to complete the remaining three form fields independently.

Figure 5: Form for test

5.4 PERFORMANCE

5.4.1 Training

During training stage, only up to two sessions were required for the participant to get accustomed to the usage of the proposed prototype. As instructions are provided through voice feedback from the Smartphone automatically, the participants only need to learn the meanings of the three tones (see Table 1) and the right response to them. Explanations were also provided on explaining how to stabilize the ruler using the Velcro strip and suggesting where to place both hands. The whole training process takes less than ten minutes per person.

5.4.2 Testing

After the training session, each participant is asked to complete the remaining three fields on their own. The test session took less than ten minutes per person. While there is little issue for most participants, it is notable that P3 required a signature template, a plastic mask with aperture to correspond

Table 2: The Feedback from Participants

Question (0=Not at all;5=Great)	P1	P2	P3	P4
Is this tool easy to learn?	5	1	2	5
Is this tool easy to use?	4	0	4	5
Is the ruler helpful in positioning the pen?	5	4	4	5
How well do you feel by using this tool compared to filling out by your own?	5	4	5	5

with standard signature area, for writing once the pen is positioned at the correct position, as P3 felt more comfortable in writing in that way. Thus, P3 ignored any feedback from the Smartphone at the writing step (the Step 6 introduced in Section 4.2) since he expected the form field should have a sufficient size to accept the signature template length.

Figure 6: Form filling result: (a) P1; (b) P2; (c) P3; (d) P4.

5.4.3 Accuracy

The accuracy of the prototype can be seen on the forms they have completed. As the participants were asked to write anything they want, most of them wrote their names as that is what they usually need to write on the form. Due to privacy concerns, only the bounding polygon is thus provided (See Figure 6 for extracted form field regions with field labels). The first form field can be ignored as it is for practice only. In the remaining three form fields, each participant is able to write them roughly inside the form field. It is also clear that P2, P3 and P4 sometimes wrote the letter before the beginning of the form field. We observed that this is because they hold the pen too low so that the color-based algorithm misinterprets the finger as the pen tip. Other than that, the positioning accuracy is quite satisfactory.

5.5 FINDINGS

A number of findings have been observed from this field trial. User evaluation and user suggestions are discussed next.

5.5.1 User Evaluation

All participants are asked to give the feedback on four questions. The questions and the feedback is consolidated in Table 2.

For the first question “Is this tool easy to learn?”, P1 and P4 gave the highest rating indicating it is very easy (5 points out of 5) to learn. P3 rated it as a bit difficult (2 points), the reason for that is that P3

is used to a physical bound box, e.g. a signature template, instead of tones as a guide to writing within a region. P2 said it was difficult to learn (1 point), citing that the clipboard is too light for him and the ruler did not slide well. In addition, as P2 cannot write, P2 has never needed to fill out the form on a piece of paper by himself, except his signature. P2 felt that it is much easier to ask someone for help as no learning is required. However, P2 understood the procedure of how to fill out a form field after learning in only one training session.

For the second question “Is this tool easy to use?”, P3 rated it very easy (5 points out of 5), as he feel no trouble in using it independently. P1 and P4 rated it easy (4 points), the reasons for that include: the precision of the positioning is set too high, they need to move the pen and ruler very slowly until it is aligned at the correct position; the Velcro strip sometimes obstructed the moving of the ruler, etc. P2 rated it very difficult (0 point). The reason for that is also about the fact that he does not write. However, we observed that he completed the form much faster than P1 and P3.

The third question is “Is the ruler helpful in positioning the pen?” The reason to ask this question is to make sure the ruler is not considered as an obstacle for the user, since it means more objects are required. Both P1 and P4 thought that it was very important (5 points), because they need the ruler to help them to position the pen. P3 rated it important (4 points). P3 thought that is a ruler is important but not ideal, as a ruler only helps to identify the lower border of the form field, not the other three borders. Similarly, P2 rated it important (4 points).

For the last question is “How well do you feel by using this tool compared to filling out by your own?”, all participants except P2 rated it great (5 points), as they could not find any other way to fill out a form by themselves. P2 rated it abit below great (4 points), the reason for that is also that P2 thought this was a good tool in assisting filling out the form independently when compared with no help, but P2 was averse to awarding a maximum value.

5.5.2 User Comments and Suggestions

In addition to the pre-designed questions, the participants have also given some valuable comments and suggestions for improving the prototype. We separate these comments into three categories: equipment, usage, and software.

Equipment

Several suggestions have been proposed by the participant to modify the current equipment used in the prototype:

- P1 suggested that the shape of the cradle base should be rectangular instead of

circular: then they can feel the orientation of the phone by only touching the cradle base.

- P1 recommend some engineering on the tool so that the ruler and the Smartphone can be moved together, as that will make sure the ruler is always observed by the Smartphone.
- P2 suggested the clipboard should be made heavier or can be fixed on the desk, so that it does not move during writing.
- P3 mentioned that the ruler is hard to move, and velcro strip is not a long lasting material.
- P3 recommend to do some engineering on the board including: 1) install tracks on both ends of the ruler; 2) add an additional ruler to locate the top boarder of the form field; 3) add additional two sliding objects to locate the left and right boarders of the form field.

Usage

P3 recommended some training on how to hold the pen, as P3 got used to putting the left hand to help to position the pen, but it blocked the ruler when using our tool. But after some learning, he feel comfortable in using only one hand to position the pen.

Software

The software itself also received some comments from the participants:

- P1 suggested the precision requirement of the software should be dropped, as a few millimeters error is not a big problem for writing a form but may potentially speed up the positioning of the ruler and pen.
- P3 and P4 commented that software for positioning in its current form works very well as the instructions are very clear.

5.6 Discussions

Further to users' feedback, we also identified a problem during the field trial. The blind people usually have their own way of holding the pen. They generally put the hand very close to the pen tip. This affects the performance of our color-based pen tip detection algorithm, as the hand and its shading could affect the correct detection of the pen tip. Therefore, as P3 suggested, training on how to hold the pen would be helpful.

During the training process, the participant were

asked to memorize three different tones. As the tones are easy to distinguish, it took less than one minute for the participant to identify each tone. In fact, we found that as long as the participant can distinguish the high pitch one (indicating stop moving the object) from others, there should be no problem to follow the form filling process as only one object is requested to move at each step.

Overall, most of the comments for improving the current prototype are focused on the equipment. This is valid because most off-the-shelf products are targeted at sighted people. Therefore, it is necessary for the equipment to be improved to satisfy their needs. In addition, making a specific equipment will greatly assist them to locate different components of the system without worrying any component being misplaced.

In terms of writing, we found that it is necessary that the blind person knows how to write. Otherwise, positioning the pen to the form field is meaningless for them. We were told that it may take about six months for a normal blind person to learn Braille. But less than fifty percent of them has learnt Braille because it is not very useful nowadays as technology advances, such as OCR and TTS. However, hardcopy forms are likely to stay with us for the foreseeable future, thus there is still a need to write on a hardcopy form. Therefore, we believe through some training, those who are blind since birth can also learn how to write in letters so that they can fill out forms independently with this prototype.

As the prototype in its current form only focuses on the positioning problem, all feedback from the user is given in relation to this navigation method. There is still more work to do, such as integrating OCR with the navigation, design feedback for more complex forms, etc.

6. CONCLUSIONS

In this paper, a prototype of form filling assistant for the blind people is presented. This prototype not only gives blind people another option to fill out a form instead of asking for a human assistant, but also introduced a dynamic fiducial marker which can help to position the pen in any part of a piece of paper with a portable camera. The prototype has been evaluated by totally blind volunteers and the results and feedback have shown that the blind users were able to complete a form independently with good precision at each form field -which is an impossible task without this prototype as all participants commented.

7. REFERENCES

[1] Y. Cho and U. Neumann. Multiring fiducial systems for scalable fiducial-tracking augmented reality.

- Presence: Teleoper. Virtual Environ.,10(6):599–612, 2001.
- [2] E. Costanza and J. Huang. Designable visual markers. In CHI '09: Proceedings of the 27th International Conference on Human Factors in Computing Systems, pages 1879–1888, New York, NY, USA, 2009. ACM.
- [3] Handel. Statistical Machine (U.S. Patent 1,915,993), 1933.
- [4] T. Ihle, H. Schirmer, and S. Fuchs. Interpretation of printed forms for blind people. In CAIP '95: Proceedings of the 6th International Conference on Computer Analysis of Images and Patterns, pages 550–555, London, UK, 1995. Springer-Verlag.
- [5] IRS. Accessible IRS tax products. <http://www.irs.gov/formspubs/>, June 2010.
- [6] H. Kato and M. Billingham. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In IWAR '99: Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality, page 85, Washington, DC, USA, 1999. IEEE Computer Society.
- [7] D. Kirk, A. Sellen, S. Taylor, N. Villar, and S. Izadi. Putting the physical into the digital: issues in designing hybrid interactive surfaces. In BCS HCI '09: Proceedings of the 2009 British Computer Society Conference on Human-Computer Interaction, pages 35–44, Swinton, UK, UK, 2009. British Computer Society.
- [8] J. Lee. Hacking the nintendo wii remote. Pervasive Computing, IEEE,7(3):39–45, july-sept.2008.
- [9] J. Marco, E. Cerezo, S. Baldassarri, E. Mazzone, and J. C. Read. Bringing tabletop technologies to kindergarten children. In BCS HCI '09: Proceedings of the 2009 British Computer Society Conference on Human-Computer Interaction, pages 103–111, Swinton, UK, UK, 2009. British Computer Society.
- [10] J.-S. Park and B.-J. Lee. Vision-based real-time camera match moving using a known marker. Optical Engineering,47(2):027201, 2008.
- [11] QuantumTechnology. Products for people who are blind. <http://www.quantumtechnology.com.au/blindness>, June 2010.
- [12] J. G. Sheridan, J. Tompkin, A. Maciel, and G. Roussos. Diy design process for interactive surfaces. In BCS HCI '09: Proceedings of the 2009 British Computer Society Conference on Human-Computer Interaction, pages 485–493, Swinton, UK, UK, 2009. British Computer Society.
- [13] A. State, G. Hirota, D. T. Chen, W. F. Garrett, and M. A. Livingston. Superior augmented reality registration by integrating landmark tracking and magnetic tracking. In SIGGRAPH, pages 429–438, 1996.
- [14] J. Varadarajan and S. Madhvanath. Digital ink to form alignment for electronic clipboard devices. In DAS '08: Proceedings of the 2008 The Eighth IAPR International Workshop on Document Analysis Systems, pages 622–628, Washington, DC, USA, 2008. IEEE Computer Society.
- [15] VisionAustralia. Taxi transport subsidy. <http://www.visionaustralia.org.au/info.aspx?page=991>, June 2010.
- [16] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg. Pose tracking from natural features on mobile phones. In ISMAR '08: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, pages 125–134, Washington, DC, USA, 2008. IEEE Computer Society.
- [17] G. Zufferey, P. Jermann, S. Do-Lenh, and P. Dillenbourg. Using augmentations as bridges from concrete to abstract representations. In BCS HCI '09: Proceedings of the 2009 British Computer Society Conference on Human-Computer Interaction, pages 130–139, Swinton, UK, UK, 2009. British Computer Society.