Supporting Everyday Activities for Persons with Visual Impairments Through Computer Vision-Augmented Touch

Leah Findlater[1], Lee Stearns[2], Ruofei Du[2], Uran Oh[2], David Ross[3], Rama Chellappa[4], Jon E. Froehlich[2]

[1]College of Information Studies, [2]Dept. of Computer Science, [4]Dept. of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742

[3]Atlanta VA R&D Center of Excellence in Vision and Neurocognitive Rehabilitation (CVNR)
1670 Clairmont Road, Decatur, GA 30033

ABSTRACT
The *HandSight* project investigates how wearable micro-cameras can be used to augment a blind or visually impaired user's sense of touch with computer vision. Our goal is to support an array of activities of daily living by sensing and feeding back non-tactile information (*e.g.*, color, printed text, patterns) about an object as it is touched. In this poster paper, we provide an overview of the project, our current proof-of-concept prototype, and a summary of findings from finger-based text reading studies. As this is an early-stage project, we also enumerate current open questions.

Categories and Subject Descriptors
K.4.2 [**Computer and Society**]: Social Issues – Assistive technologies for persons with disabilities

Keywords
Blind; visually impaired; wearable computing; computer vision; vision-augmented touch

1. INTRODUCTION
The majority of work on computer vision to support blind and visually impaired users has focused on at-a-distance information tasks such as navigation (*e.g.*, [1]) and spatial perception (*e.g.*, [3]). Instead, the *HandSight* project explores how computer vision can support proximal information accessed through touch. By sensing and feeding back non-tactile information about the physical world *as it is touched*, our goal is to support an array of activities of daily living (ADLs). This solution could, for example, allow the user to touch and read a printed page via real-time OCR and speech synthesis, touch a piece of clothing to hear suggestions for coordinating items, or pick up a bell pepper and recognize that it is red or green. Since touch is a highly attuned means of acquiring information for people with visual impairments (*e.g.*, [2]), we hypothesize that collocating sensing with the user's touch will enable new and intuitive assistive applications.

HandSight is part of a small but growing body of work on augmenting touch with wearable cameras and computer vision algorithms. Perhaps the most well-known system is *Orcam*, a recently introduced commercial head-mounted camera system designed to recognize objects and read printed text. While live demonstrations have been impressive[1], there is no academic work examining Orcam's effectiveness. This head-mounted camera approach could potentially complement HandSight's finger-mounted sensing. More closely related to HandSight are *Magic Finger* [8] and *FingerReader* [4,5]. Magic Finger uses a finger-mounted camera with the goal of "enabling any surface to act as a touch screen," although the focus is not for people with visual impairments—most problematically, the device covers the fingertip, thus interfering with tactile sensation [8]. FingerReader, in contrast, is a custom finger-mounted device with vibration motors designed to allow visually impaired wearers to read printed text by direct line-by-line scanning with the finger [4,5]. While promising, user evaluation of FingerReader to date has been limited to small qualitative studies (3-4 participants).

ASSETS '15, October 26-28, 2015, Lisbon, Portugal
ACM 978-1-4503-3400-6/15/10.
http://dx.doi.org/10.1145/2700648.2811381

(a) NanEye2C with and without 4 LEDs. (b) Ring form factor (c) Nail form factor

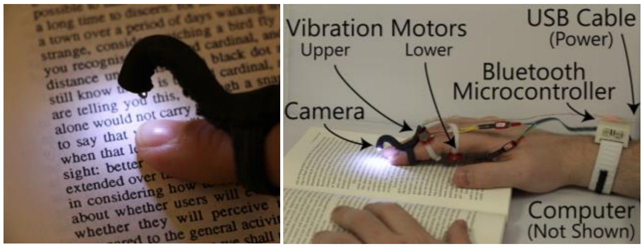
Figure 1: HandSight uses a 1×1mm^2 AWAIBA NanEye 2C camera developed for minimally invasive surgeries (endoscopies) that can capture 250×250px images at 44fps. Early form factors also shown.

This poster paper provides an overview of HandSight's hardware and software, as well as a summary of findings from our initial area of exploration—reading printed text. Because this work is at an early stage, we also cover open questions and challenges that will need to be addressed as we expand to other application areas.

2. CURRENT PROTOTYPE
Our envisioned system consists of tiny CMOS cameras (1×1mm^2) and micro-haptic actuators mounted on one or more fingers, computer vision and machine learning algorithms to support fingertip-based sensing, and a smartwatch for processing, power, and speech output. Figure 1 shows early form factors that we have been exploring, including mounting the camera on a ring or unobtrusively under an acrylic fingernail. Our overarching design goals are that HandSight should: (1) Support touch-based rather than at-a-distance interaction; (2) not hinder normal tactile function, such as covering the fingertip; (3) be easy-to-learn/use (an issue with many sensory aids) through speech and simple haptic feedback; (4) be always-available to allow for seamless transitions to/from real-world tasks; (5) be comfortable & robust to everyday stresses such as water and impact.

The current proof-of-concept prototype consists of a laptop running custom software, external speakers, a finger-mounted camera, and a custom haptic device. The camera is a self-illuminated Awaiba NanEye 2C CMOS camera with an LED ring (~40 fps, 250x250 pixels, 2.4mm diameter; Figure 1a), embedded in an adjustable ring and positioned above the finger. Because our initial application area is reading, this camera position is designed

[1] OrCam TED@NYC. YouTube Video: http://youtu.be/_3XVsCsscyw?t=4m53



(a) Finger-mounted camera **(b) Current prototype**

Figure 2: Early design for our proof-of-concept prototype with the camera placed for reading printed text. Future designs may be made much smaller, for both the camera mount and the haptic feedback.

to offer a reasonably large field of view for optical character recognition (OCR) on a printed page. Two disc vibration motors (8mm diameter, 3.4mm thick), controlled by an Arduino Pro Micro, provide haptic feedback. The motors are attached to the user's right index finger with separate Velcro rings (Figure 2b), one on top of the finger on the intermediate phalange and one below the finger on the proximal phalange.

3. INITIAL APPLICATION: READING

We have initially explored HandSight to read printed text. Compared to handheld text scanners such as mobile phone applications, a finger-based approach has the potential to mitigate camera framing issues, enable a blind reader to better understand the spatial layout of a document, and provide better control over reading pace. However, such an approach also needs to guide the reader in physically navigating a document and tracing along lines of text. While previous work has proposed audio and haptic directional finger guidance for this purpose, user studies have been limited to 3 or 4 participants [4,5,7]. The most recent of these studies underscores the need for further investigation: all 3 participants found it difficult to read text with FingerReader [5]. This provokes the question: *why*? Are finger-based cameras a viable accessibility solution for reading printed text or simply too challenging to use given issues of accurate text capture and recognition as well as finger guidance?

To further investigate the feasibility of finger-based reading, we conducted two lab studies [6]. The first study included 20 blind participants and compared audio and haptic directional finger guidance during reading tasks. Users traced their finger left to right along each line of text and the system generated text-to-speech output. In this study, we isolated the user experience from the wearable prototype's OCR performance by displaying documents on an iPad (the touchscreen provided accurate sensing of the finger location). Audio guidance played a continuous tone when the user's finger strayed off the current line, with a low pitch indicating the need for downward movement and high pitch indicating upward. For haptic guidance, the lower motor (proximal phalange) vibrated to indicate downward movement and the upper motor indicated the opposite. Participants used both types of guidance in counterbalanced order to read plain text documents and magazine-style documents (with pictures).

The two types of guidance resulted in relatively similar line tracing performance, with speeds ranging from 106–120 WPM across all conditions. Audio, however, may offer an accuracy advantage—with the magazine document participants strayed farther from the line with haptic than with audio (M=15px *vs.* 11 px away from the middle of the line; Z_{19}=-2.374, p=.018, r=.54).

Preference was almost evenly split (11 of 20 participants preferred haptic guidance), reflecting contradictions found in previous research [4,5,7]. Open-ended comments also highlighted the tradeoffs of the two types of guidance. Some participants felt that audio interfered with the speech output: *"You could focus on the audio of the text, and not be listening for other sounds"* (P8), or *"I missed a couple words because I was being distracted by the [audio]"* (P16). Those who preferred the audio reported that haptic guidance required more effort and could lead to desensitization. Finally, while several participants appreciated the direct access to layout information provided with HandSight's exploration mode, important concerns arose overall about ease of use and the amount of concentration required.

In the second study, 4 blind participants used a proof-of-concept wearable prototype (Figure 2b) and KNFB Reader iOS, a state-of-the-art OCR application, to complete reading tasks. The findings highlighted the strengths and weaknesses of the two approaches. HandSight offered additional layout information and more immediate access to text content without the need to worry about framing the complete document within the camera's field of view. But, once the document was successfully scanned, KNFB Reader offered a faster and smoother reading experience, and was preferred by all participants. Further work is needed to assess the impacts of the tradeoffs on document comprehension.

4. DISCUSSION AND OPEN QUESTIONS

The goal of the HandSight project is to support an array of activities of daily living by sensing and feeding back non-tactile information (*e.g.*, color, printed text, visual patterns) about an object as it is touched. We plan to explore several application areas beyond reading; however, many open questions remain. While a finger-based camera is always available and may mitigate camera-framing issues that can occur with mobile devices (*e.g.*, smartphones), wearables introduce practical issues of robustness and social acceptability. Future prototypes will be smaller and less obtrusive, leveraging our tiny camera form factor. We are also studying how to incorporate smartwatches more effectively, such as offloading haptic feedback to the band of the watch. Still, whether blind users are interested in wearing a finger-mounted device for accessibility is an open question.

5. ACKNOWLEDGEMENTS

This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs under Award W81XWH-14-1-0617.

6. REFERENCES

[1] Dakopoulos, D. and Bourbakis, N.G. Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey. *IEEE SMC, Part C (Applications and Reviews) 40*, 1 (2010), 25–35.

[2] Goldreich, D. and Kanics, I.M. Tactile Acuity is Enhanced in Blindness. *J. Neurosci. 23*, 8 (2003), 3439–3445.

[3] Khambadkar, V., Folmer, E. GIST: A Gestural Interface for Remote Nonvisual Spatial Perception. *Proc. ACM UIST'13*, (2013), 301–310.

[4] Shilkrot, R., Huber, J., Liu, C., Maes, P., and Chandima, N.S. FingerReader: a wearable device to support text reading on the go. *CHI '14 Extended Abstracts*, (2014), 2359–2364.

[5] Shilkrot, R., Huber, J., Wong, M.E., Maes, P., and Nanayakkara, S. FingerReader: a wearable device to explore printed text on the go. *Proc. ACM CHI'15*, (2015), 2363–2372.

[6] Stearns, L., Du, R., Oh, U., et al. Evaluating Haptic and Auditory Directional Guidance to Assist Blind Persons in Reading Printed Text Using Finger-Mounted Cameras. *In submission*.

[7] Stearns, L., Du, R., et al. The design and preliminary evaluation of a finger-mounted camera and feedback system to enable reading of printed text for the blind. *ACVR'14 Workshop at ECCV'14*, (2014).

[8] Yang, X.-D., Grossman, T., Wigdor, D., and Fitzmaurice, G. Magic finger: always-available input through finger instrumentation. Proc. ACM UIST '12, (2012), 147–156.