FingerReader: A Wearable Device to Support Text Reading on the Go

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

Citation
Roy Shilkrot, Jochen Huber, Connie Liu, Pattie Maes, and Suranga Chandima Nanayakkara. 2014. FingerReader: a wearable device to support text reading on the go. In CHI ’14 Extended Abstracts on Human Factors in Computing Systems (CHI EA ’14). ACM, New York, NY, USA, 2359-2364.

As Published
http://dx.doi.org/10.1145/2559206.2581220

Publisher
Association for Computing Machinery (ACM)

Version
Author's final manuscript

Accessed
Wed Mar 16 15:47:15 EDT 2016

Citable Link
http://hdl.handle.net/1721.1/92436

Terms of Use
Creative Commons Attribution-Noncommercial-Share Alike

Detailed Terms
http://creativecommons.org/licenses/by-nc-sa/4.0/
FingerReader: A Wearable Device to Support Text Reading on the Go

Roy Shilkrot¹
roys@media.mit.edu

Pattie Maes¹
pattie@media.mit.edu

Jochen Huber¹,²
jhuber@mit.edu

Suranga C. Nanayakkara²
suranga@sutd.edu.sg

Connie K. Liu¹
ckliu@mit.edu

¹ MIT Media Lab
75 Amherst street
Cambridge, MA 02139 USA

² Singapore University of Technology and Design
20 Dover Drive
Singapore, Singapore

Abstract
Visually impaired people report numerous difficulties with accessing printed text using existing technology, including problems with alignment, focus, accuracy, mobility and efficiency. We present a finger worn device that assists the visually impaired with effectively and efficiently reading paper-printed text. We introduce a novel, local-sequential manner for scanning text which enables reading single lines, blocks of text or skimming the text for important sections while providing real-time auditory and tactile feedback. The design is motivated by preliminary studies with visually impaired people, and it is small-scale and mobile, which enables a more manageable operation with little setup.

Author Keywords
Assistive technology; Text reading; Wearable camera; Finger worn interface;

ACM Classification Keywords
K.4.2 [Social Issues]: Assistive technologies for persons with disabilities; B.4.2 [Input/Output Devices]: Voice; I.4.8 [Scene Analysis]

Introduction
Accessing text documents is troublesome for visually impaired (VI) people in many scenarios, such as reading
text on the go and accessing text in less than ideal conditions (i.e. low lighting, columned text, unique page orientations, etc.) Interviews we conducted with VI users revealed that available technologies, such as screen readers, desktop scanners, smartphone applications, eBook readers, and embossers, are commonly under-utilized due to slow processing speeds or poor accuracy. Technological barriers inhibit VI people’s abilities to gain more independence, a characteristic widely identified as important by our interviewees.

In this paper, we present our work towards creating a wearable device that could overcome some issues that current technologies pose to VI users. The contribution is twofold:

- First, we present results of focus group sessions with VI users that uncovered salient problems with current text reading solutions and the users’ ideographs of future assistive devices and their capabilities. The results serve as grounds for our design choices.

- Second, we present the concept of local-sequential text scanning, where the user scans the text progressively with the finger, which presents an alternative solution for problems found in existing methods for the VI to read printed text. Through continuous auditory and tactile feedback, our device allows for non-linear reading, such as skimming or skipping to different parts of the text, without visual guidance. To demonstrate the validity of our design we conducted early user studies with VI users to assess its real-world feasibility as an assistive reading device.

Related Work
Giving VI people the ability to read printed text has been a topic of keen interest in academia and industry for the better part of the last century. The earliest attainable evidence of an assistive text-reading device for the blind is the Optophone from 1914 [6], however the more notable effort from the mid 20th century is the Optacon [10], a steerable miniature camera that controls a tactile display. In table 1 we present a comparison of recent methods for text-reading for the VI based on key features: adaptation for less-than-perfect imaging, target text, UI tailored for the VI and method of evaluation. We found a general presumption that the goal is to consume an entire block of text at once, while our approach focuses on local text and gives the option of skimming over the text as well as reading it thoroughly. We also handle non-planar and non-uniformly lit surfaces gracefully for the same reason of locality, and provide truly real-time feedback.

Prior work presents much of the background on finger worn devices for general public use [12, 15], although in this paper we focus on a wearable reading device for the VI.

Assistive mobile text reading products
Academic effort is scarcely the only work in this space of assistive technology, with end-user products readily available. As smartphones became today’s personal devices, the VI adopted them, among other things, as assistive text-reading devices with applications such as the kNFB kReader [1], Blindsight’s Text Detective [5]. Naturally, specialized devices yet exist, such as ABiSee’s EyePal ROL [3], however interestingly, the scene of wearable assistive devices is rapidly growing, with OrCam’s assistive eyeglasses [2] leading the charge.
Table 1: Recent efforts in academia of text-reading solutions for the VI. Accuracy is in precision (P) recall (R) form, as reported by the authors.

User Needs Study
To guide our work, we conducted two focus group sessions ($N_1 = 3$, $N_2 = 4$, all of them congenitally blind users) to gain insights into the users’ text reading habits and identify issues with current technologies. Moreover, we introduced early, non-functional prototypes of the FingerReader as stimuli to both get the participants’ opinion on the form factor and elicit potential design recommendations. Both sessions took place over a course of 5 hours on average, therefore we summarize only the most relevant findings:

- All of the participants used a combination of flatbed scanners and mobile devices (e.g., a camera-equipped smartphone) to access printed text in their daily routine.

- Flatbed scanners were considered straightforward; most problems occur when scanning prints that do not fit on the scanner glass. Mobile devices were favored due to their portability, but focusing the camera on the print was still considered tedious. However, the participants considered both approaches inefficient. As one participant put it: “I want to be as efficient as a sighted person”.

- Primary usability issues were associated with text alignment, word recognition accuracy, processing speed of OCR software, and unclear photography due to low lighting issues. Slow return of information was also an issue, as the overall time to digitize a letter-sized page was estimated to be about 3 minutes.

- Participants were interested in being able to read fragmented text such as a menu, text on a screen, or a business card, and warped text on canned goods labels. They also preferred the device to be small in order to allow hands-free or single-handed operation.

Taking this information into consideration, we decided to design a mobile device that alleviates some common issues and enables a few requested features: capable of skimming, works in real time, and provides feedback using multiple modalities.
FingerReader: A wearable reading device

FingerReader is an index-finger wearable device that supports the VI in reading printed text by scanning with the finger (see Figure 1c). The design continues the work we have done on the EyeRing [12], however this work features novel hardware and software that includes haptic response, video-processing algorithms and different output modalities. The finger-worn design helps focus the camera at a fixed distance and utilizes the sense of touch when scanning the surface. Additionally, the device does not have many buttons or parts in order to provide a simple interface for users and easily orient the device.

Hardware details

The FingerReader hardware expands on the EyeRing by adding multimodal feedback via vibration motors, a new dual-material case design and a high-resolution mini video camera. Two vibration motors are embedded on the top and bottom of the ring to provide haptic feedback on which direction the user should move the camera via distinctive signals. The dual material design provides flexibility to the ring’s fit as well as helps dampen the vibrations and reduce confusion for the user (Fig. 1b). Early tests showed that users preferred signals with different patterns, e.g. pulsing, rather than vibrating different motors, because they are easier to tell apart.

Software details

To accompany the hardware, we developed a software stack that includes a text extraction algorithm, hardware control driver, integration layer with Tesseract OCR [17] and Flite Text-to-Speech (TTS) [4], currently in a standalone PC application.

The text-extraction algorithm expects an input of a close-up view of printed text (see Fig 2). We start with image binarization and selective contour extraction. Thereafter we look for text lines by fitting lines to triplets of pruned contours; we then prune for lines with feasible slopes. We look for supporting contours to the candidate lines based on distance from the line and then eliminate duplicates using a 2D histogram of slope and intercept. Lastly, we refine line equations based on their supporting contours. We extract words from characters along the selected text line and send them to the OCR engine. Words with high confidence are retained and tracked as the user scans the line. For tracking we use template matching, utilizing image patches of the words, which we accumulate with each frame. We record the motion of the user to predict where the word patches might appear next in order to use a smaller search region. Please refer to the code\(^1\) for complete details.

When the user veers from the scan line, we trigger a tactile and auditory feedback. When the system cannot find more word blocks along the line we trigger an event to let users know they reached the end of the printed line. New high-confidence words incur an event and invoke the TTS engine to utter the word aloud. When skimming, users hear one or two words that are currently under their finger and can decide whether to keep reading or move to another area.

Our software runs on Mac and Windows machines, and the source code is available to download\(^1\). We focused on runtime efficiency, and typical frame processing time on our machine is within 20ms, which is suitable for realtime processing. Low running time is important to support randomly skimming text as well as for feedback, for the user gets an immediate response once a text region is detected.

---

\(^1\) Source code is currently hosted at: [http://github.com/royshil/SequentialTextReading](http://github.com/royshil/SequentialTextReading)
Evaluation
We evaluated FingerReader in a two-step process: an evaluation of FingerReader’s text-extraction accuracy and a user feedback session for the actual FingerReader prototype from four VI users. We measured the accuracy of the text extraction algorithm in optimal conditions at 93.9% ($\sigma = 0.037$), in terms of character misrecognition, on a dataset of test videos with known ground truth, which tells us that part of the system is working properly.

User Feedback
We conducted a qualitative evaluation of FingerReader with 4 congenitally blind users. The goals were (1) to explore potential usability issues with the design and (2) to gain insight on the various feedback modes (audio, haptic, or both). The two types of haptic feedbacks were: fade, which indicated deviation from the line by gradually increasing the vibration strength, and regular, which vibrated in the direction of the line (up or down) if a certain threshold was passed. Participants were introduced to FingerReader and given a tablet with text displayed to test the different feedback conditions. Each single-user session lasted 1 hour on average and we used semi-structured interviews and observation as data gathering methods.

Each participant was asked to trace through three lines of text using the feedbacks as guidance, and report their preference and impressions of the device. The results showed that all participants preferred a haptic fade compared to other cues and appreciated that the fade could also provide information on the level of deviation from the text line. Additionally, a haptic response provided the advantage of continuous feedback, whereas audio was fragmented. One user reported that “when [the audio] stops talking, you don’t know if it’s actually the correct spot because there’s no continuous updates, so the vibration guides me much better.” Overall, the users reported that they could envision the FingerReader helping them fulfill everyday tasks, explore and collect more information about their surroundings, and interact with their environment in a novel way.

Discussion and Summary
We contribute a novel concept for text reading for the VI of a local-sequential scan, which enables continuous feedback and non-linear text skimming. It is implemented in a novel tracking-based algorithm that extracts text from a close-up camera view and a finger-wearable device.
FingerReader presents a new way for VI people to read printed text locally and sequentially rather than in blocks like existing technologies dictate. The design is motivated by a user needs study that shows the benefit in using continuous multimodal feedback for text scanning, which again shows in a qualitative analysis we performed. We plan to hold a formal user study with VI users that will contribute an in-depth evaluation of FingerReader.

Acknowledgements
We wish to thank the Fluid Interfaces and Augmented Senses research groups, K. Ran and J. Steimle for their help. We also thank the VIBUG group and all VI testers.

References
[1] KNFB kReader mobile, 2010. http://www.knfbreader.com/products-kreader-mobile.php.
[2] OrCam, 2013. http://www.orcam.com/.
[3] ABiSee. EyePal ROL, 2013. http://www.abisee.com/products/eye-pal-rol.html.
[4] Black, A. W., and Lenzo, K. A. Flite: a small fast run-time synthesis engine. In ITRW on Speech Synthesis (2001).
[5] Blindsight. Text detective, 2013. http://blindsight.com/textdetective/.
[6] d’Albe, E. F. On a type-reading optophone. Proc. of the Royal Society of London. Series A 90, 619 (1914), 373–375.
[7] Ezaki, N., Bulacu, M., and Schomaker, L. Text detection from natural scene images: towards a system for visually impaired persons. In ICPR (2004).
[8] Hanif, S. M., and Prevost, L. Texture based text detection in natural scene images—a help to blind and visually impaired persons. In CVHI (2007).
[9] Kane, S. K., Frey, B., and Wobbrock, J. O. Access lens: a gesture-based screen reader for real-world documents. In Proc. of CHI, ACM (2013), 347–350.
[10] Linvill, J. G., and Bliss, J. C. A direct translation reading aid for the blind. Proc. of the IEEE 54, 1 (1966).
[11] Mattar, M. A., Hanson, A. R., and Learned-Miller, E. G. Sign classification for the visually impaired. UMass-Amherst Technical Report 5, 14 (2005).
[12] Nanayakkara, S., Shilkrot, R., Yeo, K. P., and Maes, P. EyeRing: a finger-worn input device for seamless interactions with our surroundings. In Augmented Human (2013).
[13] Pazio, M., Niedzwiecki, M., Kowalik, R., and Lebiedz, J. Text detection system for the blind. In EUSIPCO (2007), 272–276.
[14] Peters, J.-P., Thillou, C., and Ferreira, S. Embedded reading device for blind people: a user-centered design. In ISIT (2004).
[15] Rissanen, M. J., Vu, S., Fernando, O. N. N., Pang, N., and Foo, S. Subtle, natural and socially acceptable interaction techniques for ringfaces: Finger-ring shaped user interfaces. In Distributed, Ambient, and Pervasive Interactions. Springer, 2013, 52—61.
[16] Shen, H., and Coughlan, J. M. Towards a real-time system for finding and reading signs for visually impaired users. In Computers Helping People with Special Needs. Springer, 2012, 41–47.
[17] Smith, R. An overview of the tesseract OCR engine. In ICDAR (2007), 629–633.
[18] Yi, C., and Tian, Y. Assistive text reading from complex background for blind persons. In Camera-Based Document Analysis and Recognition. Springer, 2012, 15–28.