
EyeRing: A Finger-worn Assistant

Suranga Nanayakkara^{1, 2}

suranga@sutd.edu.sg

Roy Shilkrot¹

roys@media.mit.edu

Patricia Maes¹

pattie@media.mit.edu

¹ MIT Media Laboratory

75 Amherst Street

Cambridge, MA, 02142

² Singapore University of Technology and Design

20 Dover Drive

Singapore, 138682

Abstract

Finger-worn interfaces are a vastly unexplored space for interaction design. It opens a world of possibilities for solving day-to-day problems, for visually impaired people and sighted people. In this work we present EyeRing, a novel design and concept of a finger-worn device. We show how the proposed system may serve for numerous applications for visually impaired people such as recognizing currency notes and navigating, as well as helping sighted people to tour an unknown city or intuitively translate signage. The ring apparatus is autonomous, however it is counter parted by a mobile phone or computation device to which it connects wirelessly, and an earpiece for information retrieval. Finally, we will discuss how finger worn sensors may be extended and applied to other domains.

Author Keywords

Pointing-based Interaction; Wearable Assistive Device; Intuitive Interfaces

ACM Classification Keywords

H.5.2 User Interfaces; H.5.1 Multimedia Information Systems; I.4.8 Scene Analysis; K.4.2 Social Issues: Assistive technologies for persons with disabilities

General Terms

Algorithms; Design; Human Factors

Copyright is held by the author/owner(s).
CHI'12, May 5–10, 2012, Austin, Texas, USA.
ACM 978-1-4503-1016-1/12/05.



(a) Symbol's WSS 1000



(b) Optical Finger Mouse



(c) The Optacon II

**Figure 1:** Some related work.

Introduction

Despite the attention finger-worn interaction devices have received over the years, there is still much room for innovative design. Earlier explorations of finger-worn interaction devices (some examples are shown in Figure 1) may be divided into a few subspaces according to how they are operated: Pointing [1], [2]; Tapping/Touching [3–5]; Gesturing [3], [6–8]; Pressing/Clicking On-Device [9].

Firstly, we wish to make the distinction between pointing gestures with the finger touching the object and pointing in free air. Our system is based on performing Free-air Pointing (FP) gestures, as well as Touch Pointing (TP) gestures. TP gestures utilize the natural touch sense, however the action trigger is not based on touch sensitivity of the surface, rather on an external sensor. Pointing devices based on TP gestures, as a reading aid for the blind date back to the *Optophone* and later the *Optacon*¹. However, the rise of cheap and small photo-sensory equipment, such as cameras, revolutionized the way low-vision people read or interact with visual interfaces. Recently Chi et al presented *Seeing With Your Hand* [10], a glove apparatus that uses TP gestures. Other assistive devices that are using imaging technology but not TP gestures are Primpo's iSONIC² and the i-Cane³ which act both as a white cane and as a visual assistant that can tell the ambient lighting condition and colors of objects. The haptic element of TP gestures is interesting especially in the case of assistive technologies for the visually impaired. This enables

them to get additional feedback on the object they want to interact with.

FP gestures on the other hand, are rooted in human behavior and natural gestural language. This was shown to be true by examining gestural languages of different cultures [11]. Usually FP gestures are used for showing a place or a thing in space - a passive action. However, augmenting FP for information retrieval is an interesting extension. Previous academic work in the field of FP gestures, revolved around control [2] and information retrieval [1]. These works and others utilize a specialized sensor, usually an infrared connection, between the pointing finger and the target. This implies the environment to be rigged especially for such interaction. We chose to use a generalized approach by using a general-purpose camera. This choice breaks the bonds of dimensionality of a single signal source or sensor, as well as the bonds of wavelength as it operates in the wider, visible spectrum.

The desire to replace an impaired human visual sense or to augment a healthy one had a strong influence on the design and rationale behind EyeRing. Most of the work around FP and some TP gestures (e.g. the Optical Finger Mouse) are aimed towards sighted people. At the initial stage of this project, we chose to focus on a more compelling aspect of exploring how visually impaired people may benefit from finger-worn devices.

In this paper, we describe the EyeRing prototype, a few applications of EyeRing for visually impaired people and some future possibilities. Finally we discuss our plans of extending this work beyond the assistive interfaces domain.

¹ <http://en.wikipedia.org/wiki/Optacon>

² <http://primpo.com/eng/products/isonic.html>

³ <http://www.i-cane.nl/en/home>

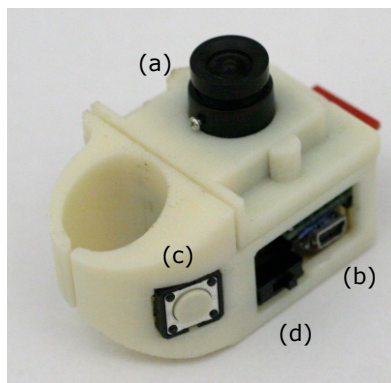


Figure 3: EyeRing prototype. The case was 3D printed with ABS nylon, and the electronics fit completely inside the skeletal compartment. (a) VGA mini camera, (b) Mini-USB port for recharging and reprogramming, (c) Trigger button, (d) On/Off switch.

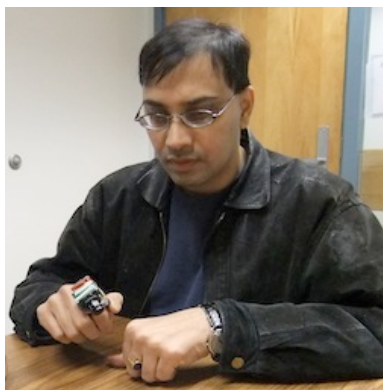


Figure 4: A blind user wearing EyeRing System. We are in the stages of developing a more slick design, which will have a center of gravity closer to the finger and overall smaller size.

EyeRing – A Finger Worn Assistant

Our proposed system is composed of a finger-worn device with an embedded camera, a computation element embodied as a mobile phone, and an earpiece for information loopback. The finger-worn device is autonomous and wireless, and includes a single button to initiate the interaction. Information from the device is transferred to the computing element where it is processed, and the results are transmitted to the headset for the user to hear. Overview of the EyeRing system is shown in (Figure 2). The current implementation of finger-worn device uses a TTL Serial JPEG Camera, 16 MHz AVR processor, a bluetooth module, 3.7V polymer Lithium-ion battery, 3.3V regulator, and a push button switch. These components are packaged into a wearable ABS nylon casing (Figure 3). Speech processing engine and computer vision engine were implemented on a mobile phone running Android 2.2.

A user needs to pair the finger-worn device with the mobile phone application only once and henceforth a Bluetooth connection will be automatically established when both are running. A visually impaired user indicated this as an essential feature. Typically, a user would single click the pushbutton switch on the side of the ring using his thumb (Figure 4). At that moment, a snapshot is taken from the camera and the image is transferred via bluetooth to the mobile phone. An Android application on the mobile phone then analyzes the image using our computer vision engine. Type of analysis and response depends on the preset mode (color, distance, currency, etc.). Upon analyzing the image data, the Android application uses a Text-to-Speech module to read out the information through a headset. Users may change the preset mode by double-

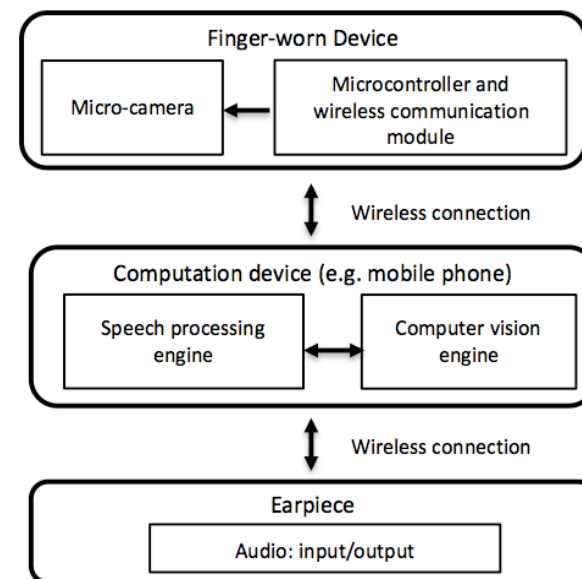


Figure 2: Overview of EyeRing system

clicking the pushbutton and giving the system brief verbal commands such as “distance”, “color”, “currency”, etc., which are subsequently recognized.

Applications For The Visually Impaired

The task of replacing the optical and nervous system of the human visual sense is an enormous undertaking. Thus, we choose to concentrate on learning the possible interaction mechanics for three specific scenarios outlined in this section.

1. Virtual Walking Cane: Compared to a steel cane, a finger-worn device used for navigation is certainly less obtrusive, as well as fashionable and appealing. The essence of this application is to provide an approximate estimation of the clear walking space in front of the



(a) The user takes two photos pointing at the ground in front.



(b) A sample pair of photos



(c) A rendering of the reconstructed 3D scene.



(d) The estimated floor model.

Figure 5: Virtual Walking Cane. Estimating free walking space using a single EyeRing.

holder of the ring. Users must use FP gestures to take pictures of the space in front of them, by pointing the camera and clicking, with some motion between the shots. In a continuous-shooting mode (video), which is currently not supported in our prototype, there is no need for repeated clicking. The system clearly notifies the approximate free space in front. For this application, we employ the concept of Structure from Motion (SfM). Upon receiving the two images, an algorithm to recover the depth is performed. The general outline of the algorithm is as follows: (a) the two images are scanned for salient feature points and then affine-invariant descriptors are calculated for them, (b) the features in both images are matched into pairs, (c) the pairs are used to get a dense distance map from the location of the camera in the first photo [12], (d) we use a robust method to fit a model of a floor to the data, and return the distance of the clear walking path in front. By repeatedly taking photos with motion, equivalent of moving a steel cane, we check the recovered 3D mapping of the floor and objects for any obstacles in the way of the user. Figure 5 outlines the above process.

2. Currency Detector: This application is intended to help the user to identify currency bills (1\$, 5\$, 10\$, 20\$, 100\$) to aid with payments. The interaction process is simple; a user would simply point index finger to a currency note (using a TP gesture then moving the finger back) and click the button. The system will voice out what the note is (Figure 6). A detection algorithm based on a Bag of Visual Words (BoVW) approach [13] scans the image and makes a decision on the type of note it sees. We use Opponent Space [14] SURF features to retain color information, for notes detection. Our vocabulary was trained to be of

1000 features long, and we use a 1-vs.-all SVM approach for classifying the types of notes. For training we used a dataset of 800 images using k-fold cross-validation, and 100 images withheld for testing. As of writing these lines the overall recognition rate is over 80% with a 0.775 kappa statistic.

3. Color Detector: This application of EyeRing aids a visually impaired person to understand the color of an object. Again, the user interaction is simple; users simply touch point (TP gesture) to an object and click the button on the ring to deliver an image for processing (Figure 7). The system analyses the image and returns the average color via audio feedback. We use a calibration step to help the system adjust to different lighting conditions. A sheet of paper with various colored boxes is printed, and a picture of it is taken. We rectify the region in the image so that it aligns with the colored boxes, and then extract a sample of the pixels covering each box. For predictions we use a normal distribution, set to the maximum likelihood of the perceived color.

User Reactions

We consulted a visually impaired user to get some initial feedback on our prototype design. This user was born sighted, but he lost his eyesight due to Retinitis Pigmentosa (genetic incurable disease) around 2000. He uses a white cane to get around and relies on speech technology for reading and using a computer. He also uses an iPod touch (with voice-over option), which has applications for detecting color and currency. He was given the EyeRing system and explained how he could use it. His first reaction was “*this is neat*”. All three applications were very relevant to him. He commented that his white cane often falls down when



Figure 6: Identifying a currency note



Figure 7: Identifying a color

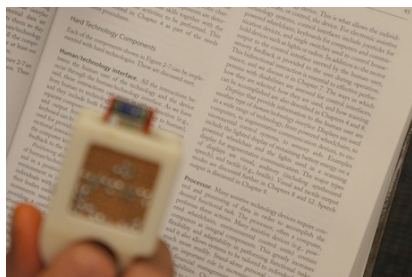


Figure 8: Concept of using EyeRing to read printed text.

he leaves it by his chair and he has to find it on his own or ask someone else to pick it up for him. He prefers the idea of having a ring on his finger that helps him to navigate. He wishes if we can develop the system in to a visual depth scanner, where he can wave his finger to get depth information. He also added that he spent a lot of time at the beginning of the day trying to pick a shirt (with a specific color). He often has to make payments with cash. For both color and currency identification he has to find his iPod touch, unlock screen, browse to the application, open it, and take a picture to get a result. With EyeRing, he said, "*[It] can't get any easier*". He is often unsure about the "right distance" to hold his iPod when taking a picture for color or currency detection. He commented that TP gestures (i.e. touching the object with his finger tip) are going to be "*very helpful*". He also added that we could improve the color detection and currency detection algorithms. Overall, he was satisfied with the system as the interactions are seamless and he could use a single device to do three things by changing the mode. This single-user-study gave us some useful insights and we plan to validate his comments with a formal user study.

Future Applications

EyeRing is still very much a work in progress. Future applications using EyeRing rely on more advanced capabilities of the device, such as real-time video feed from the camera, higher computational power, and additional sensors like gyroscopes and a microphone. These capabilities are currently in development for the next prototype of EyeRing.

Reading Non-Braille: Visually impaired people are mostly bound to reading Braille or listening to audio

books. However the amount of written material that are in Braille or audio transcribed is limited. The fact remains that natural interaction in our world requires decent visual abilities for reading. This application will allow the visually impaired to read regular printed material using EyeRing. The user simply touches the printed surface with the tip of the finger (TP gesture) and moves along the lines. Naturally, a blind person cannot see the direction of the written lines. For that reason we plan to implement an algorithm to detect the misalignment between the movement of the finger and the direction of the text, correcting the user's movement using the audio feedback. Figure 8 illustrates the concept of this application.

Tourist Helper: We plan to extend EyeRing applications to domains beyond assistive technology. For example, tourists visiting a new city often rely on maps and landmarks for navigation. Recently, locationing systems, inertial sensors and compasses in mobile devices, and readily available dense mapping of most major cities, replaced the usage of paper maps. However, even with augmented reality applications such as Layar⁴, the UI for on-foot navigation is still cumbersome. Our proposed application relies on a much more natural gesture, such as simply pointing at the wanted landmark, asking "What is that?" and hearing its name.

Conclusion and Future Work

EyeRing suggests a novel interaction method for both visually impaired and sighted people. We choose to base the interaction on a human gesture that is ubiquitous in any language and culture – pointing with

⁴ <http://www.layar.com/>

the index finger. This has determined the nature and design of the ring apparatus, location of the camera and trigger. The applications we presented for EyeRing emerge from the current design. Preliminary feedback received from a visually impaired user supports that EyeRing assistive applications are intuitive and seamless. We are in the process of conducting a more formal and rigorous study to validate this. One of the biggest challenges is creating the supporting software that works in unison with this unique design. However, we believe that adding more hardware such as a microphone, an infrared light source or a laser module, a second camera, a depth sensor or inertial sensors, will open up a multitude of new uses for this specific wearable design.

Acknowledgements

We would like to thank the Fluid Interfaces group for their insightful remarks, Amit Zoran for the design of the prototype casing, and Savithru Jayasinghe for his work on the Android application development.

References

- [1] Merrill, D. and Maes, P. Augmenting looking, pointing and reaching gestures to enhance the searching and browsing of physical objects. *In Proc. Pervasive'07*, (2007) 1–18.
- [2] Tsukada, K. and Yasumura, M. Ubi-finger: Gesture input device for mobile use. *In Proc. APCHI'02*, 1 (2002), 388–400.
- [3] Lee, J., Lim, S. H., Yoo, J. W., Park, K. W., Choi, H. J. and Park, K. H. A ubiquitous fashionable computer with an i-Throw device on a location-based service environment. *In Proc. AINAW'07*, 2 (2007), 59–65.
- [4] Fukumoto, M. and Suenaga, Y. 'FingerRing': a full-time wearable interface. *In Proc. CHI'94*, ACM Press (1994), 81–82.
- [5] Fukumoto M. A finger-ring shaped wearable handset based on bone conduction. *In Proc. ISWC '05*, (2005), 10–13.
- [6] Rekimoto, J. Gesturewrist and gesturepad: Unobtrusive wearable interaction devices. *In Proc. ISWC '01*, (2001), 21–27.
- [7] Zhai, S., Milgram, P. and Buxton, W. The influence of muscle groups on performance of multiple degree-of-freedom input. *In Proc. CHI'96*, ACM Press (1996), 308–315.
- [8] Zimmerman, T.G., Lanier, J., Blanchard, C., Bryson, S and Harvill, Y. A hand gesture interface device. *In Proc. CHI '87*, ACM Press (1987), 189–192.
- [9] Chatterjee R. and Matsuno, F. Design of A Touch Sensor Based Single Finger Operated Wearable User-Interface Terminal. *In Proc. SICE-ICASE'06*, (2006), 4142–4147.
- [10] Chi, L. Y., Ryskamp, R. A., Gomez, L. R. P., Ho, H. and Brin, S. Seeing with your hand. Google Patents, 2011.
- [11] McNeill, D. *Language and gesture*, Cambridge University Press, 2000.
- [12] Hartley, R. and Zisserman, A. *Multiple view geometry in computer vision*, Cambridge University Press, 2000.
- [13] Csurka, G., Dance, C., Fan, L., Willamowski, J. and Bray, C. Visual categorization with bags of keypoints. *In Proc. ECCV'04*, (2004), 59–74.
- [14] Van de Sande, K., Gevers, T. and Snoek, C.. Evaluation of color descriptors for object and scene recognition. *In Proc. CVPR'08*, (2008), 1–8.