Camera to Assist People with Visual Impairments while Shopping

ROGER BOLDU, Augmented Human Lab, The University of Auckland, New Zealand

ALEXANDRU DANCU, Augmented Human Lab, Singapore University of Technology and Design, Singapore

DENYS J.C. MATTHIES, Augmented Human Lab, The University of Auckland, New Zealand

THISUM BUDDHIKA, Augmented Human Lab, The University of Auckland, New Zealand

SHAMANE SIRIWARDHANA, Augmented Human Lab, The University of Auckland, New Zealand

SURANGA NANAYAKKARA, Augmented Human Lab, The University of Auckland, New Zealand

# Introduction (*Heading 1*)

People with Visual Impairments (PVI) experience greater difficulties with daily tasks, such as supermarket shopping. Identifying and purchasing an item proves challenging for PVI.

. In contrast with FingerReader, FingerReader2.0 incorporates a highly integrated hardware design, as it is standalone, wearable, and not tethered to a computer. Software-wise, the prototype utilizes a deep learning system, relying on a hybrid, an on-board and a cloud-based model. The advanced design significantly extends the range of mobile assistive technology, particularly for shopping purposes

# Description

The contribution of this successor paper is threefold:

(1) Technical development of a standalone, wearable finger-worn prototype, FingerReader2.0, in the form of a small ring with a camera and touch input, which is able to recognize products and notes to help PVI acquire products.

(2) Introduction of the user-centered design process to understand the needs of our user group, as well as discovering challenges and opportunities for the design and evaluation of an assistive finger-worn smart eye.

(3) Compiled insights for designing wearable assistive pointing interfaces for PVI based on interviews, focus groups, and a field study using the FingerReader2.0 prototype inside a supermarket.

# System ArchitectureC:\Users\gelud\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Снимокsa.png

The prototype contains three hardware components:

(1) a ring with an embedded camera and a touch interface (see Figure 2 - B). The location of the camera enables the system to capture the image of what the user is pointing at, while simultaneously allowing the user to control the device through the touch interface. The ring is tethered to the

second component,

(2) a wristband that contains the processing unit. This processing unit is composed of a system on board, a wireless module (Wifi+BLE), and a battery (see Figure 2 - C). The processing unit transmits the captured images to a third component,

(3) a smartphone through Wifi communication. The smartphone performs the image analysis and delivers the information to a user through a Bluetooth Headset or through the phones speaker.

Hardware description:

## *1 Ring Hardware Prototype*

Electronics:

The ring incorporates a VGA camera module ov7675. This camera is connected to a main custom made PCB. On the left side of the ring, there is a 15mm x 20mm touch interface, implemented on an external custom made PCB. This sends the results of the touch input to the external processing unit.

## *Wristband Hardware Prototype*

Electronics: The wristband processing unit is based on a custom made PCB that operates on an Intel Edison SOM (system on module). This wristband also includes a Dual-core Intel Atom 500MHz processor, 1GB DDR3 RAM, 4GB eMMC flash, Bluetooth 4.0, Wifi, Wi-Fi Direct. The system runs an embedded Linux Yocto 1.1. The power consumption tests show that the device can last approximately 3.5 hours.

## *Phone Application*

A hotspot network is opened by the phone to connect with the wristband. Once this communication occurs, an Android phone application receives and analyzes the images from the wristband. The received image is analyzed with an on-board identification and an external scene description. Once the image analysis result is returned in plain text, the system synthesizes the audio by using Google text-to-speech engine. The user receives the output through the phone speaker or via Bluetooth headphones.

## *Image Understanding and Communication*

In order to identify the object the user is pointing at, developers evaluated various state-of-the-art image recognition libraries using our hardware prototype. In conclusion, the creators decided to implement a hybrid approach, which analyzes the captured image in parallel (1) using an on-board deep learning algorithm, as well as (2) sending it to an external cloud vision API.

# 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.

Prior work presents much of the background on ﬁnger

worn devices for general public use [12,15], although in

this paper the focus is 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.

Also article “Digital Digits: A Comprehensive Survey of Finger Augmentation Devices”[18] strives to survey the entire body of work on finger augmentation devices. New finger-augmenting devices (FADs) appear annually in major academic venues of the HCI community, in the consumer market as new products, and in prominent news and popular media outlets. Finger-augmenting devices may be defined as finger-worn devices with an additional augmentation other than their form that provide a supplemental capability for one or more fingers using the finger itself as a central element.

There are seven generic form factors used in FADs, which are also illustrated in Figure 4: rings, distal addendums, whole-finger addendums, fingernail addendums, sleeves, thumb addendums, and, finally, components mounted on the palm that support the FADs.

On the other hands similar hardware design is used for developing new solutions for other purposes. The articles below have similar design but are used for completely different purposes.

The article “Thumb-In-Motion: Evaluating Thumb-to-Ring Microgestures for Athletic Activity”[16] is evaluated a one-handed thumb-to-ring gesture interface to quickly access information without interfering with physical activity, such as running. By a pilot study, the most minimal gesture set was selected, particularly those that could be executed reflexively to minimize distraction and cognitive load. The evaluation revealed that among the selected gestures, the tap, swipe-down, and swipe-left were the most ‘easy to use’. 

Interestingly, motion does not have a significant effect on the ease of use or on the execution time. However, interacting in motion was subjectively rated as more demanding. Finally, the gesture set was evaluated in real-world applications, while the user performed a running exercise and simultaneously controlled a lap timer, a distance counter, and a music player.

The article MoveSpace: On-body Athletic Interaction for Running and Cycling describe a method of interaction involving tapping specific locations on the body, identify candidate locations for running and cycling, and compare them in a series of controlled experiments with athletes. 

A purpose-built prototype measures speed of interaction and gives feedback cues for athletes to report the physical effects on the activity itself. Our results suggest that specific locations are faster and have minimal disruption to movement, even under induced fatigue conditions. The overall method is fast - 1.31s for running and 1.65s for cycling. Preferred locations differ significantly across sports, with stable body parts ranking higher. We effectively demonstrated the use of a single hand for interaction during running with two distinct tap gestures. A set of guidelines inform the design of new sports technologies.

# 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.

[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), 373375.

[10] Linvill, J. G., and Bliss, J. C. A direct translation

reading aid for the blind. Proc. of the IEEE 54, 1

(1966).

[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).

[15] Rissanen, M. J., Vu, S., Fernando, O. N. N., Pang,

N., and Foo, S. Subtle, natural and socially

acceptable interaction techniques for ringterfaces:

Finger-ring shaped user interfaces. In Distributed,

Ambient, and Pervasive Interactions. Springer, 2013,

52—61

[16]<http://alexandrudancu.com/pdf/Thumb-in-Motion_SUI2018.pdf>

[17]http://alexandrudancu.com/wp-content/uploads/2016/05/movespace.pdf

[18] http://ahlab.org/sites/default/files/a30-shilkrot\_0.pdf