

# Write-it-Yourself with the Aid of Smartwatches: A Wizard-of-Oz Experiment with Blind People

Syed Masum Billah

Vikas Ashok

IV Ramakrishnan

Department of Computer Science, Stony Brook University, NY, USA

{sbillah, vganjiguntea, ram}@cs.stonybrook.edu

## ABSTRACT

Working with non-digital, standard printed materials has always been a challenge for blind people, especially writing. Blind people very often depend on others to fill out printed forms, write checks, sign receipts and documents. Extant assistive technologies for working with printed material have exclusively focused on reading, with little to no support for writing. Also, these technologies employ special-purpose hardware that are usually worn on fingers, making them unsuitable for writing. In this paper, we explore the idea of using off-the-shelf smartwatches (paired with smartphones) to assist blind people in both reading and writing paper forms including checks and receipts. Towards this, we performed a Wizard-of-Oz evaluation of different smartwatch-based interfaces that provide user-customized audio-haptic feedback in real-time, to guide blind users to different form fields, narrate the field labels, and help them write straight while filling out these fields. Finally, we report the findings of this study including the technical challenges and user expectations that can potentially inform the design of *Write-it-Yourself* aids based on smartwatches.

## Author Keywords

Accessibility; visual impairments; blind; writing aid; wearables; directional guidance; audio-haptic; smartwatch.

## INTRODUCTION

Although the Web has become the dominant medium for doing many transactional tasks such as online banking, online form-based registrations, and online shopping, printed materials such as paper forms, postal mails, checks, contracts and shopping credit-card receipts still abound in our daily lives. While paper-based transactions, especially filling out paper forms and affixing signatures on printed materials is straightforward for sighted people, they present significant barriers to blind people, notably writing. People with vision impairments form a sizeable population—over 21 million

people with significant vision loss in the US alone and over 39 million blind people worldwide [29].

To assist blind people with printed materials, existing assistive-technology (AT) solutions primarily employ OCR via cameras embedded in smartphones (KNFB Reader [3], Seeing AI [19]), Head-Mounted Displays [21] and custom-designed hardware (FingerReader [25]). All these ATs focus exclusively on reading printed content, and provide little-to-no support for writing (e.g. writing checks and signing credit-card receipts). Besides, a bigger problem is that in some cases an AT solution, especially ones based on special-purpose hardware designed to be worn on fingers [25], interfere with writing.

In this paper, we report on the feasibility of using off-the-shelf smartwatches with built-in audio-haptic feedback, as a *write-it-yourself* (WiY) AT to aid blind users to work with paper forms and paper receipts *independently*. To this end, we conducted a Wizard-of-Oz (WoZ) study with 12 blind participants to capture and understand the technical challenges, and user requirements and expectations that can potentially inform the design of smartwatch-based WiY aids. Our choice of smartwatches for the study was dictated by its growing popularity as the preferred wearable [23].

## RELATED WORK

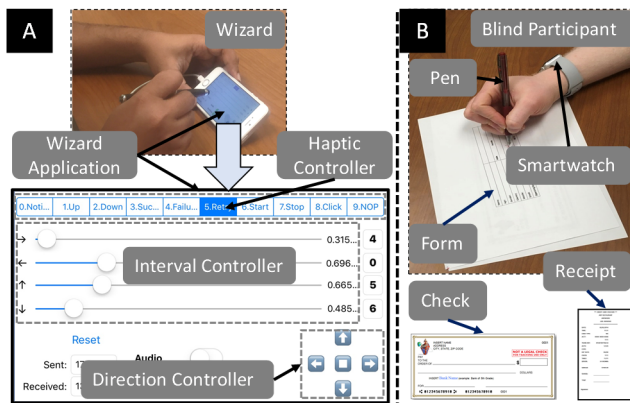
The use of braille, one of the oldest tactile-based read/write systems for the blind, has been steadily declining [20], hastened by advances in digital technologies, particularly speech, computer vision and OCR. A number of stand-alone reading applications, leveraging these advances, have emerged—examples include the pioneering Kurzweil Scanner [4] for desktops; KNFB Reader [3], Seeing AI [19] and Text Detective [5] for mobile phones; and FingerReader [25], OrCam [21], HandSight [26] for wearables devices. OCR, that underpin all of these stand-alone reading applications have several limitations—they are unusable in poor lighting conditions; they require careful camera framing so that a target object is completely visible and centered within the camera's field of view [15, 18]; they do not support complex documents and spatial data [16], and cannot determine which blocks of text to read, and in what order [9]. Reading apps on wearables such as [21, 25, 26] address some of these limitations. But they have other problems—they require specialized hardware and cameras and the ones designed to be worn on fingers interfere with writing. These reasons have prevented them from gaining mainstream acceptance.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org). *IUI'18, March 7–11, 2018, Tokyo, Japan*

© 2018 Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-4945-1/18/03...\$15.00

<https://doi.org/10.1145/3172944.3173005>



**Figure 1. (A) The Wizard interface with customization option; (B) A blind participant writing on printed materials (e.g., a form, check) with a regular pen while wearing a smartwatch.**

Crowd-sourcing of object recognition in images, using, say Mechanical Turk, has also been explored as an alternative to OCR, notable examples being VizWiz [8] and BeMyEyes [1]. Crowd-sourced solutions can conceivably be used as a WiY aid. But two major problems with these solutions are that the blind users lose privacy and have to depend on others to help with writing.

Audio-haptic cues for way-finding and for exploration of maps have been used in numerous projects [6, 11, 17, 18, 22, 24, 27]. These works use some combination of synthesized speech and sonification and haptic patterns for effective discrimination [10, 12, 14]. The drawback with these audio-haptic approaches is that they require extraneous hardware including several specialized haptic motors, ranging from 4 to 8, and mounted camera sensors. Pairing a smartphone with an off-the-shelf smartwatch that comes with built-in audio-haptic feedback is an attractive approach to solving the WiY problem without the need for such extraneous hardware, and thus has the potential to become an unobtrusive viable mainstream writing aid for blind individuals. The feasibility of such an approach is reported in the sequel.

## METHOD

The experimenter in our WoZ study (i.e., the wizard) controlled the direction of participants' hand movements with audio-haptic feedback provided by a smartwatch worn by the participants (see Figure 1). The choice of the audio-haptic cues for different directions however, was self-customized by the participants themselves.

## Participants

We recruited 12 blind participants (labeled as P1 to P12) through local mailing lists and word-of-mouth. No participant had any hearing or speech impairment. They varied in age from 30 to 60 (mean=40.5, median=35, SD=11.1), and gender (7 males, 5 females). All participants reported prior experience with audio-haptic feedback, which they had gained from using smartphones, game controllers, gloves and other devices. All participants except P5, P10 and P11, stated that writing was very much a part of their daily routine, e.g., paying bills, teaching their children, shopping,

or at work. Those 3 participants only knew how to write their name and signature—2 of them (P5 and P11) were born blind, and P10 who lost his vision 20 years back had mostly forgotten how to write. All participants said that they sought help from others when writing on paper.

## Apparatus

The smartwatch interface was designed as an Apple Watch App running in tandem with a *companion* Wizard App on an iPhone (see Figure 1); the latter let the experimenter transmit audio-haptic feedback instructions in real-time to the Watch App to control the hand movements of the participants during the writing process. In other words, the Wizard App was set up to help the wizard guide blind participants to different *fields* (e.g., name, signature, tip) on the printed forms by continuously transmitting these audio-haptic feedback instructions. The printed materials used in the study were sample templates of a bank check, restaurant receipt, and a personal-information form as shown in Figure 1.B.

## Wizard App

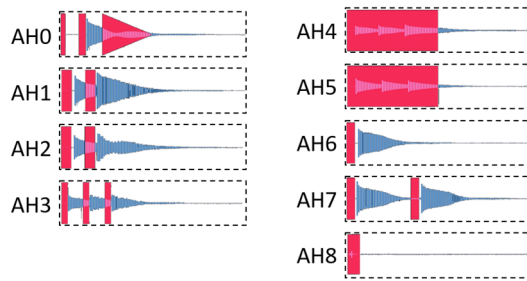
The Wizard App (see Figure 1.A) has arrow keys (indicated by “Direction Controller” in the Figure 1A) that allows the experimenter to send different audio-haptic signals for different directions of hand movement. The exact audio-haptic pattern for each direction can be configured using the “Haptic Controller”. Further, the time-intervals for the selected patterns can also be adjusted by moving the 4 sliders shown in the “Interval Controller”.

## Audio-Haptic Feedback

We used the 9 default audio-haptic patterns in Apple watch [2]. Figure 2 shows these patterns (labeled as AH0 to AH8). Each pattern consists of one or more individual haptic and audio signals. The amplitude and the duration of the haptic signals are shown in pink, while the audio signals are depicted in blue. For any hand-movement direction, one single pattern was relayed continuously, with the interval between two such patterns were set based on preference.

## Procedure

The experimenter started the study with a brief background survey, and then introduced the participants to the Watch App. Prior to the study, participants were first exposed to the different audio-haptic signals, and then allowed to select the ones they preferred (including tuning of the associated vibration-stimulus intervals) for distinguishing between different directional hand movements. The participants were then asked to complete the following three tasks (ordering was counterbalanced): (A) Fill out the details in a sample check; (B) Write the tip amount, total amount and sign a sample restaurant receipt; and (C) Fill out the name, email, and country fields in a sample paper form. Participants unfamiliar with writing could sign the form fields instead of writing-in actual data.



**Figure 2. Nine audio-haptic feedback patterns (AH0 to AH8), pink for haptic and blue for audio signals.**

For each task, the participants began by first taking a picture of the printed material with an iPhone provided by the experimenter. To take a picture, the participants were instructed to hold the iPhone directly above and parallel to the sample paper on the table, and continuously reposition the camera according to synthesized audio instructions, until the camera could capture the entire form within the picture frame. This was done to create an illusion that the participants were interacting with a fully-functional WiY application, rather than a wizard. To further strengthen this illusion, the participants were explicitly told that a background application sent this picture to the WiY application on the Watch via Bluetooth for processing and automatically generating audio-haptic guidance instructions.

After taking the picture of the paper form, the participants proceeded with the task by first placing their pen-tip at its top-left corner. They then moved their hand in the direction indicated by the received audio-haptic signal. The wizard ensured that the participants were guided to different fields one-by-one via their preferred audio and haptic cues. Feedback was also provided to ensure that the participants maintained a straight line while writing out the required information in the various printed fields.

At the end, the experimenter conducted a semi-constructed interview with preselected topics such as participant's writing proficiency, strategies, and usability of WiY. The interview culminated with participants making suggestions and recommendations. The entire session lasted for 2 hours.

### Measures

All tasks were video-recorded. Audio-haptic preferences of users, including vibration-stimulus intervals were logged. Other collected data included task-completion times, user accuracy in correctly interpreting haptic cues, and task success rate. A task was considered a *success* if the required information was written reasonably straight and acceptably close to the intended fields; the judgement was left to the discretion of the experimenter. In total, the participants wrote into 108 fields (12 \* 3 tasks/participant \* 3 fields/task).

### RESULTS

**Participants' preferences:** The participants' preferences of audio-haptic patterns for the four directions of hand movement, as well as the interval specification for each of these preferred patterns are shown in Figure 3. The order of

	AH0	AH1	AH2	AH3	AH4	AH5	AH6	AH7	AH8	
P1	R	U				L		D		
P2	R					U	D		L	
P3	L				R	U	D			
P4		R	L			U	D			
P5	R				U				D	
P6		R		L	U				D	
P7	R			D	U	L				
P8				R	L			U	D	
P9	L				R			U	D	
P10		L		U	D		R			
P11	R				L		D	U		
P12	R	L		U			D			

Range (sec.)	Color
[0.3, 0.5]	
[0.5, 0.8]	
[0.8, 1.0]	
[1.0, 1.3]	
[1.3, 1.5]	

**Figure 3. Participants' preferences (rows) for audio-haptic patterns (columns) for right (R), left (L), up (U), and down (D) directions. Color shades denote interval ranges for patterns.**

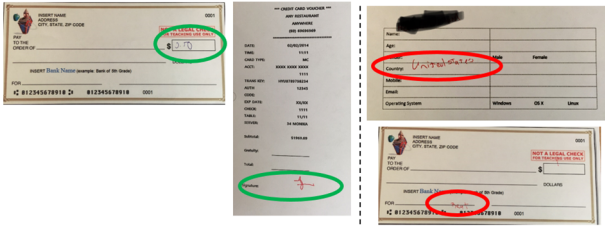
preference was as follows: AH5 (18%), AH0 (16%), AH6, AH7 (both 12%), AH5, AH1 (both 10%); AH3 (8%), AH8 (6%) and AH2 (4%). Again, see Figure 2 for each pattern.

The participants stated that they could easily recognize the long persistent haptic pattern in AH5, as well as the long but gradually decaying haptic pattern in AH0. The gradually decaying nature of AH0 prompted them to set short intervals (0.3s to 0.8s), whereas, the persistent nature of AH5 made them to select a relatively longer interval (0.5s to 1.3s).

Since AH4 and AH5 are almost similar with the only difference being the underlying audio tone, majority (10) of the participants selected only one of these patterns during customization. Only participants P3 and P7 indicated that they were comfortable distinguishing AH4 from AH5 using audio cues, and so they picked both these patterns.

Interestingly, as seen in Figure 3, a couple of participants (P8 and P9) used a single pattern (AH7) to represent multiple directional movements (Up and Down), by selecting different intervals for the same pattern. P8 also used a single pattern (AH4) with different intervals for representing Right and Left directional movements. When inquired, these participants stated that they did not want to memorize too many patterns. Some participants selected certain patterns because these patterns reminded them of certain events in their daily lives. For example, P2, P3, P4 choose AH5 for Up because they felt that the tone of AH5 sounded like their morning wakeup alarm. Similarly, P10 stated that AH1 and AH6 resembled the indicator sounds of a car, and P12 said that AH1 sounded like heartbeat to her. We also noticed a trend in selecting patterns and intervals for the opposite direction movements. For example, if a strong pulse with a short interval was chosen for a direction, then a weaker pulse with a long interval was chosen for the opposite direction.

**Task completion time:** The average completion times for the 3 tasks (A, B, and C) were 185s (SD: 15.5s), 163s. (SD: 10.3s), and 240s (SD: 19.6s) respectively. To compare these completion times with that of sighted people, one author and 6 MS students (sighted) performed the same tasks. On an average, they took 33.2s (SD: 6.1s), 12.5s (SD: 5.7s) and 25.9s (SD: 8.3s) for tasks A, B, and C respectively.



**Figure 4.** Ovals are target fields on the papers. Successful (left of dotted line) and unsuccessful (to the right) writing samples.

**Interpretation accuracy:** The participants interpreted the received audio-haptic patterns with an average accuracy of 97%. Participants themselves stated that since they handpicked their own patterns, it was easy for them to interpret the directional instructions.

**Task success rate:** The success rates as defined earlier, were 84%, 68%, and 45% for tasks A, B, and C respectively. Figure 4 presents examples of successful and unsuccessful tasks. The primary reason for failures was lack of prior calibration to accommodate hand-to-pen distance. Also, different participants had different pen-holding styles, which made it a little difficult for the wizard to predict where exactly on paper they would end up writing. For example, in Figure 4 (right), even if the participant P4 was properly guided to a field on the check, the written entry ended up overlapping with the line, thereby resulting in a failure. Another reason for failures was the limited space allotted to certain fields, especially in Tasks B and C. This was further exacerbated by the communication overhead between the Wizard app and the Watch app; in some cases, the participants accidentally navigated past the target location before receiving a ‘stop’ haptic feedback.

### User Requirements and Expectations

Feedback from participants shaped the formulation of these user expectations and interface design requirements:

**U1. Goal notification:** A distinct audio-haptic pattern to notify users that they have reached the destination field.

**U2. Dimension notification:** Audio feedback indicating the dimensions of a selected field in a relatable manner, such as “*this field has the width of a pencil*”, or “*length is less than a credit card*”.

**U3. Discrete intervals:** The selection of vibration intervals on a nominal scale with categories like “slow”, “medium”, “fast”, and “very fast”, instead of a continuous scale.

**U4. Intelligent companion smartphone app:** Narrate the text under fingertips like VoiceOver while navigating the picture of the printed material taken by the smartphone app. Also, double tapping on any form label in the picture should trigger their smartwatch to guide them to the corresponding field on the actual printed material.

**U5. Picture taking with the smartwatch:** Allow taking pictures of the printed materials with a smartwatch as an alternative for blind people who are not adept at taking pictures with smartphone apps such as KNFB Reader.

**U6. Augmenting with a ruler or a credit card:** Allow using a ruler or a credit card to assist with writing along a straight line, in addition to audio-haptic feedback.

**U7. Wrist-Pace control:** Control the speed of wrist movements with audio cues such as “slow down” and “speed up”. Towards this, increasing or decreasing the haptic-pattern interval was not recommended since it was hard to distinguish different patterns with short intervals.

### Discussion

The study revealed the following technical challenges:

**C1. High precision motion tracking:** Tracking algorithm should have millimeter range precision since users move their hands over a Letter size page.

**C2. Detecting when the user starts writing:** Tracking algorithm needs to know when the pen-tip touches the paper to provide audio-haptic guidance to help users maintain a straight line while writing.

**C3. Calibration:** Differences in the size of hands and pen-holding styles require sophisticated wrist-to-pen calibration to accurately track pen-tip location.

**C4. Communication-overhead related issues:** Proactively prevent the user from accidentally navigating past the target location on the form due to the unavoidable communication overhead between the smartphone and smartwatch.

**C5. Mapping between picture and print material:** Techniques are needed to estimate dimensional details of the printed material by analyzing the corresponding picture.

While most of these challenges are open problems, related research works in robotics can potentially be applied to our WiY system to address some of them. For example, for C1, user can initially register the top-left corner as the origin by pressing the crown of the watch when the pen tip is at that corner and further tracking can be made relative to the origin, by measuring short-range displacements using motion sensor data [13]. For C2 and C3, ideas proposed in [7, 28] where information (e.g., typed word) is inferred from smartwatch sensor data, can be used. For C4, the ideas underlying Proportion Integral and Derivative (PID) Controllers can be utilized. Finally, for C5, recent developments in the fields of computer vision and augmented reality can be explored.

### CONCLUSION

Write-it-Yourself aids using smartwatches have the potential to become a transformative assistive technology that can empower blind people to work with printed materials independently without having to depend upon others; hence can make them more employable. Our WoZ study with blind participants gathered critical requirements and identified technical challenges that need to be addressed for realizing WiY aids in practice. This is a topic for future research.

### ACKNOWLEDGMENTS

We thank the anonymous reviewers for their insightful feedback. This research is supported by NSF: IIS-1447549, NEI/NIH: R01EY02662, NIDILRR: 90IF0117-01-00.

## REFERENCES

1. Be My Eyes. Retrieved from <http://bemyeyes.com/>.
2. 2017. Haptic Feedback. Retrieved from <https://developer.apple.com/watchos/human-interface-guidelines/user-interaction/haptic-feedback/>.
3. Knfb Reader. Retrieved from <http://www.knfbreader.com/>.
4. 1996. Kurtzweil 1000 for Windows. Retrieved from <https://www.kurzweilededu.com/products/kurzweil-1000-v14-windows.html>.
5. Text Detective. Retrieved from <https://itunes.apple.com/us/app/text-detective/id541494875?mt=8>.
6. Bardot, S., Brock, A., Serrano, M., and Jouffrais, C., 2014. Quick-Glance and in-Depth Exploration of a Tabletop Map for Visually Impaired People. In *Proceedings of the 26th Conference on l'Interaction Homme-Machine*. ACM, 165-170.
7. Beltramelli, T. and Risi, S., Deep-Spying: Spying Using Smartwatch and Deep Learning. Retrieved from <https://arxiv.org/abs/1512.05616>.
8. Bigham, J. P., Jayant, C., Ji, H., Little, G., Miller, A., Miller, R. C., Tatarowicz, A., White, B., White, S., and Yeh, T., 2010. Vizwiz: Nearly Real-Time Answers to Visual Questions. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*. ACM, 1-2.
9. Borodin, Y., Bigham, J. P., Dausch, G., and Ramakrishnan, I. V., 2010. More Than Meets the Eye: A Survey of Screen-Reader Browsing Strategies. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*. ACM, 1-10.
10. Brewster, S. and Brown, L. M., 2004. Tactons: Structured Tactile Messages for Non-Visual Information Display. In *Proceedings of the Fifth Australasian User Interface Conference (AUIC2004)*. ACS, 15-23.
11. Brock, A., Kammoun, S., Macé, M., and Jouffrais, C., 2014. Using Wrist Vibrations to Guide Hand Movement and Whole Body Navigation. *i-com* 13, 3, 19-28.
12. Choi, S. and Kuchenbecker, K. J., 2013. Vibrotactile Display: Perception, Technology, and Applications. *Proceedings of the IEEE* 101, 9, 2093-2104.
13. ChRobotics, Using Accelerometers to Estimate Position and Velocity. Retrieved from <http://www.chrobotics.com/library/accel-position-velocity>.
14. Hong, J., Stearns, L., Froehlich, J., Ross, D., and Findlater, L., 2016. Evaluating Angular Accuracy of Wrist-Based Haptic Directional Guidance for Hand Movement. In *Proceedings of the 42nd Graphics Interface Conference*. Canadian Human-Computer Communications Society, 195-200.
15. Jayant, C., Ji, H., White, S., and Bigham, J. P., 2011. Supporting Blind Photography. In *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 203-210.
16. Kane, S. K., Frey, B., and Wobbrock, J. O., 2013. Access Lens: A Gesture-Based Screen Reader for Real-World Documents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 347-350.
17. Lawrence, M. M., Martinelli, N., and Nehmer, R., 2009. A Haptic Soundscape Map of the University of Oregon. *Journal of Maps* 5, 1, 19-29.
18. Manduchi, R. and Coughlan, J. M., 2014. The Last Meter: Blind Visual Guidance to a Target. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 3113-3122.
19. Microsoft, 2017. Seeing Ai. Retrieved from <https://www.microsoft.com/en-us/seeing-ai/>.
20. NFB, 2016. Statistical Facts About Blindness in the United States.
21. OrCam, 2017. OrCam: See for Yourself. Retrieved from <https://www.orcam.com/myeye/>.
22. Pielot, M., Poppinga, B., Heuten, W., and Boll, S., 2011. A Tactile Compass for Eyes-Free Pedestrian Navigation. In *Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part II*. Springer-Verlag, 640-656.
23. Piwek, L., Ellis, D. A., Andrews, S., and Joinson, A., 2016. The Rise of Consumer Health Wearables: Promises and Barriers. *PLoS Medicine* 13, 2 (02/02), e1001953.
24. Rümelin, S., Rukzio, E., and Hardy, R., 2011. Naviradar: A Novel Tactile Information Display for Pedestrian Navigation. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 293-302.
25. Shilkrot, R., Huber, J., Ee, W. M., Maes, P., and Nanayakkara, S. C., 2015. Fingerreader: A Wearable Device to Explore Printed Text on the Go. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2363-2372.
26. Stearns, L., Du, R., Oh, U., Jou, C., Findlater, L., Ross, D. A., and Froehlich, J. E., 2016. Evaluating Haptic and Auditory Directional Guidance to Assist Blind People in Reading Printed Text Using Finger-Mounted Cameras. *ACM Trans. Access. Comput.* 9, 1, 1-38.
27. Tsukada, K. and Yasumura, M., 2004. Activebelt: Belt-Type Wearable Tactile Display for Directional Navigation. In *UbiComp 2004: Ubiquitous Computing: 6th International Conference, Nottingham, Uk, September 7-10, 2004. Proceedings*, N. DAVIES, E.D. MYNATT and I. SIIO Eds. Springer Berlin Heidelberg, Berlin, Heidelberg, 384-399.
28. Wang, H., Lai, T. T.-T., and Choudhury, R. R., 2015. Mole: Motion Leaks through Smartwatch Sensors. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 155-166.
29. WHO, 2014. World Health Organization. Visual Impairment and Blindness. Retrieved from <http://www.who.int/mediacentre/factsheets/fs282/en/>.