A Wearable Face Recognition System Built into a Smartwatch and the Visually Impaired User

Laurindo de Sousa Britto Neto[1,2], Vanessa Regina Margareth Lima Maike[2], Fernando Luiz Koch[3], Maria Cecília Calani Baranauskas[2], Anderson de Rezende Rocha[2] and Siome Klein Goldenstein[2]
[1]*Department of Computing, Federal University of Piauí (UFPI), Teresina, Brazil*
[2]*Institute of Computing, University of Campinas (UNICAMP), Campinas, Brazil*
[3]*Samsung Research Institute, Campinas, Brazil*

Keywords: Human-computer Interaction, Assistive Technology, Computer Vision, Accessibility, Wearable Device.

Abstract: Practitioners usually expect that real-time computer vision systems such as face recognition systems will require hardware components with high processing power. In this paper, we present a concept to show that it is technically possible to develop a simple real-time face recognition system in a wearable device with low processing power – in this case an assistive device for the visually impaired. Our platform of choice here is the first generation Samsung Galaxy Gear smartwatch. Running solely in the watch, without pairing to a phone or tablet, the system detects a face in the image captured by the camera, and then performs face recognition (on a limited dictionary), emitting an audio feedback that either identifies the recognized person or indicates that s/he is unknown. For the face recognition approach we use a variation of the K-NN algorithm which accomplished the task with high accuracy rates. This paper presents the proposed system and preliminary results on its evaluation.

1 INTRODUCTION

In 2013, the World Health Organization estimated that 285 million people worldwide have visual disabilities, of which 39 million are blind and 246 have low vision (W. H. Organization, 2013). Daily tasks such as walking, reading and recognizing objects or people may be very difficult or even impossible for those who are blind or have low vision. Technology can assist the visually impaired in some of these tasks, providing them more autonomy and social inclusion. In particular, the field of Computer Vision has a lot to contribute to Assistive Technologies (Manduchi and Coughlan, 2012), since, in a way, it allows a machine to replace the user's lost sight. In this paper, we focus on the twofold challenge of running a facial recognition in a wearable device to assist visually impaired users in recognizing people who are in their surroundings. One part of the challenge lies in the technological aspects of the proposal, and the other part lies in the social-technical aspects, i.e., the interaction between the user, the technology and everything else in the context of use.

For instance, imagine a scenario in which a visually impaired person walks into an environment where silence and discretion are required, such as a work meeting or a library. Under usual circumstances she would have to disrupt the silence to know who are the other people present in the environment. However, with the use of a face recognition system embedded into a wearable device, the user could accomplish the task with the required discretion. For this to be possible, it would be necessary, on the technological end, to have efficient facial recognition algorithms installed into a hardware that has compatible processing power and that is small enough to be wearable. On the social-technical end, the feedbacks provided by the system to the user would have to be easily understandable, efficient and discrete; the camera present in the device could not invade the privacy of the people surrounding the user or make them uncomfortable; finally, the way in which the user would wear the devices could not cause embarrassment.

The described system may seem impossible to

Britto Neto L., Maike V., Koch F., Baranauskas M., Rocha A. and Goldenstein S..
A Wearable Face Recognition System Built into a Smartwatch and the Visually Impaired User.
DOI: 10.5220/0005370200050012
In *Proceedings of the 17th International Conference on Enterprise Information Systems* (ICEIS-2015), pages 5-12
ISBN: 978-989-758-098-7

accomplish, but in the next few sections, we will present a proof of concept that shows how it is technically possible to develop a simple, yet quite effective, real-time face recognition system, running in a wearable device with low processing power. We will also present initial user tests that show the interaction between people and the proposed system, investigating the potential gains users can have from the system. The wearable platform we use here is the first generation Samsung Galaxy Gear smartwatch. As this model is not the newest, it has less powerful hardware than the later ones, and it is assumed that if the system works well in the limited device, it should work better in the more advanced ones. The Galaxy Gear wristwatch features a 1.9 Megapixel camera on the wrist band, which is good enough for the system we propose. Additionally, having the camera attached to the wrist allows the smartwatch to be used in hands-free operations.

Our prototype uses a library of known subjects that need to be registered prior to recognition – we do not use Internet or social-media searches to find potential matches. The smartwatch constantly acquires images, analyzes them in search of a person's face, and then gives audio feedback of that analysis. In the case of an unknown face, the system allows the registration of a new instance of an existing person, or of a new individual. Since the first generation of the Galaxy Gear runs the Android OS, the system also ran even better on a Samsung Galaxy Note 3 smartphone.

This paper is organized as follows: Section 2 describes the literature in the face recognition area focusing on wearable devices in the aid of the visually impaired, with a variety of different approaches; Section 3 describes the Samsung Galaxy Gear smartwatch; Section 4 describes the developed system; Section 5 describes the dataset used and the experiment performed as a preliminary evaluation of the system; and Section 6 concludes this work and points out further work.

2 RELATED WORK

We have performed a search on digital libraries looking for papers that approach the problem of using wearable devices to aid the visually impaired. In this section we present an overview of the works we found, in order to characterize the current state of the art of the problem we are trying to solve.

Pun et al. (2007) present a survey on assistive devices for sight-handicapped people. The survey covers works that use video processing for converting visual data into an alternative rendering modality, such as auditory or haptic. Most of these studies focuses on daily tasks such as navigation and object detection, but not on people recognition.

We can see an extensive literature review on face recognition for biometrics in Tistarelli and Grosso (2010) and Zhao et al. (2003) – the literature focusing on accessibility is more scarce. Krishna et al. (2005) developed a pair of sunglasses with a pinhole camera, which uses the Principal Component Analysis (PCA) algorithm (Kistler and Wightman, 1992) for face recognition. The idea is to be able to later evolve the system from face to emotion, gesture and facial expressions recognition. The sunglasses system was validated with a highly controlled dataset, which uses a precisely calibrated mechanism to provide robust face recognition.

Kramer et al. (2010) present a smartphone that provides audible feedback whenever a face from a database enters or exits the scene. Their detection algorithm runs in a server that uses the VeriLook face technology (NEUROtechnology, 2014). In contrast, in our system, the face recognition algorithms are running within the wearable device itself.

Astler et al. (2011) used a camera atop a standard white cane to perform face recognition using the Luxand FaceSDK (Luxand, 2013), and to identify six kinds of facials expressions using the Seeing Machines FaceAPI[1].

Tanveer et al. (2012) developed a system called FEPS, which uses Constrained Local Model algorithm for facial expressions recognition providing audible feedback, and Fusco et al. (2012) proposed a method which combines face matching and identity verification modules in feedback.

As we see in the survey of Pun et al. (2007), there are several studies conducted to create more assistive devices for the blind and low-vision people. Few reports are presented on systems that make use of smartwatch. The first is the FreevoxTouch (FreevoxTouch, 2014), a smartwatch created for the visually impaired that runs on an Android platform. Currently, it has the following functions: speaking watch, memorecorder, music player and a stopwatch/countdown. The smartwatch is entirely controlled through a touch screen, and all clock functions can be set to have an audio feedback.

Porzi et al. (2013) developed a gesture recognition system for a smartwatch that increases its usability and accessibility to assist people with

[1] http://www.seeingmachines.com/

visual disabilities. The user presses the smartwatch's display to start the gesture input. Then, the user performs a gesture and the signals generated by the smartwatch's integrated accelerometers are sent via Bluetooth to a smartphone. These signals are processed and then the system recognizes the gesture and activates the corresponding function. When the task is completed, the user receives vibration feedback. Moreover, the system has two modules: one for identifying wet floor signs and one for automatic recognition of predefined logos. A downside of it is that the smartwatch cannot be directly programmed.

Watanabe et al. (2014) proposed an activity and context recognition method in which the user carries a neck-worn receiver comprising a microphone, and small speakers on his wrists that generate ultrasounds. The system uses the volume of the received sound and the Doppler effect to recognize gestures. The system recognizes the place where the user is in and the nearby people by ID signals generated by speakers placed in rooms and on people. The authors presented the device and considered that the proposed method can be used with the Samsung Galaxy Gear smartwatch.

2.1 Face Recognition

In order to succeed, real face-recognition systems have to perform, really well, a series of complex tasks. Usually they have to detect faces, normalize them, extract descriptors, and then perform the recognition. Not all steps are present in every system, and in some methods the extraction of descriptors and the face-recognition are done together.

The most commonly used face detector is the presented by Viola and Jones (2004). Introduced first in the 2001 Conference on Computer Vision and Pattern Recognition CVPR, it presents a real-time robust algorithm for face detection and face tracking that uses Haar functions, integral images, and boosting on weak classifiers, ultimately offering efficiency and requiring less computational complexity.

Dalal and Trigg (2005) developed a descriptor named Histogram of Oriented Gradients (HOG), used to describe characteristics of objects of interest based on image gradients and borders. Other descriptors that use spatio-temporal information are the Local Binary Pattern (LBP) (Ahonen et al., 2006) and its variations, such as the Volume Local Binary Pattern (VLBP), by Zhao and Pietikainen (2007), and the Extended VLBP (EVLBP), by Hadid et al. (2007).

There are several classic face recognition methods, such as the Eigenfaces (Turk and Pentland, 1991) and the Fisherfaces (Belhumeur et al., 1997) based in PCA. They were not used in our proposal because they would add complexity to the processes of adding new people to the database and of determining the distance threshold for recognition. An initial analysis showed that the trade-off between this complexity and the possible performance gains did not pay off.

Li et al. (2013) proposed a complex framework that used a multi-modal sparse coding approach to utilize Depth information for face recognition. Other approaches using infrared images (Chen et al., 2003; Wilder et al., 1996) and 3D depth maps (Gordon, 1991) were also explored to achieve face recognition. Research about the possibility of analysing face images by modelling local facial features (Wiscott et al., 1997) were performed.

3 SAMSUNG GALAXY GEAR

The Samsung Galaxy Gear (GEAR) is a smart device shaped wristwatch (smartwatch) equipped with a 800 MHz processor, 512MB RAM, 4GB internal memory, the Android 4.2.2 operating system, two microphones, a speaker, Bluetooth and a 1.9 Megapixel camera on the wristband. It was developed to be used together with the Samsung Galaxy Note 3 smartphone. Thus, the user can make calls or other tasks of the smartphone through the smartwatch. The two devices communicate by Bluetooth, and every audio feedback can be heard through a stereo Bluetooth headset.

This wearable device comes with the Samsung S Voice application installed, a software that allows the user to perform voice-operated tasks, such as dialing a phone number, sending a text message, opening an app, and playing music, all from the smartwatch. Therefore, the S-Voice can be used to aid the visually impaired.

Moreover, the GEAR has accelerometer and gyroscope sensors, making possible the use of a gesture recognition system like in Porzi et al. (2013). This is especially useful in situations where the interaction through voice commands may not be used (such as during a meeting), or when they may not work properly (such as crowded scenarios or noisy environments).

4 SYSTEM OVERVIEW

The system we developed was named Gear Face Recognition (GFR). First, the user must open the app. There are two ways of doing this: through the S Voice application, or by setting a shortcut to open the application. In the first case, it is necessary to run S Voice by pressing the smartwatch's physical power/home button twice, and then giving the voice command associated to the app. In the second case, the user simply touches the top of the watch's display and slides it down. When the GFR opens, an audio feedback indicates that the app is running.

Our prototype system uses the camera of the GEAR to perceive the user's surroundings. As soon as a face is detected, an audio feedback is given, indicating that a person's face is being framed by the camera. In this moment, the user and the camera have to stand still for a few seconds, to finish the framing. Next, the system performs the face recognition and provides an audio feedback that characterizes the identified person, such as a ringtone, a sound, or a voice recording. Subjects must be previously registered in the system for the face recognition and a different audio can be associated to each person. Unknown subjects are mapped to a common audio feedback.

Our face detection module is based on the sample code provided by the OpenCV4Android[2] library. We extract the rectangular image of the detected face in video frames.

To run on the watch's limited hardware, we use the K-NN algorithm (Cover and Hart, 1967) with 3,780-dimensional HOG descriptors for the face recognition approach. Figure 1 illustrates this conversion. The value of hyperparameter K can be set according to the amount of registered samples per person. Initially, as a default value we use $K = 1$, as we have only few samples per person.

HOG descriptors have shown good results to represent features set for face identification (Schwartz et al., 2012). Moreover, HOG has a controllable degree of invariance to local geometric transformations, providing invariance to translations and rotations smaller than the local spatial or orientation bin size (Dalal and Triggs, 2005).

To improve the accuracy of the K-NN, we used temporal coherence over the video's sequential frames (sliding window) – we classify each frame within the temporal sliding window, and the most voted person is the final classification.

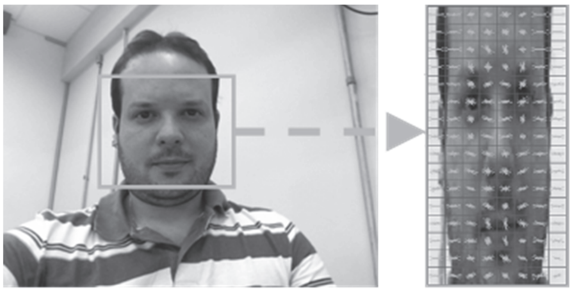

Figure 1: Example of image conversion in HOG descriptor.

A person may be classified as unknown when the unknown person class wins the voting. A vote is computed for the unknown class when the distance from the sample to all the nearest neighbours is greater than a threshold distance. The threshold was set empirically based on observations of the distance values. The rational of this decision is that distances between samples from the same person tend to be smaller than the distances between samples from different people. The value for the threshold distance may vary depending on the camera resolution. The higher the quality and resolution of images captured by the camera, the smaller the threshold distance value. A more formal analysis shows that this hypothesis assumes that the classes are separable by a plane in the HOG high-dimensional space.

We created a prototype with a simple interface for user interaction (Figure 2). When the system detects an unknown face (unknown sample), we can add this sample to a new person or to an already registered person, simply by touching the smartwatch's display to capture the face's rectangle. If a new person is being registered, then the system asks to record an audio to associate with that person: touch the display to start recording and we touch it again to finish.

If an already registered person is not recognized by the system, we create the possibility of adding new samples to an already registered person. This serves to increase the robustness of the face recognition performed by the K-NN by adding new samples of the same person to the dataset, increasing the variability of the data for the same person. From the description, it is possible to note that the registration interface is not yet ready for visually impaired users. However, studies to improve the feedback of the registration interface are being conducted so that it can also be used by people with visual disabilities.

[2] http://opencv.org/platforms/android.html

Figure 2: Gear Face Recognition: an unknown person (left), adding a sample (center) and a recognized person (right).

5 EXPERIMENT SETUP AND PRELIMINARY RESULTS

A pilot experiment was conducted with the intent of finding out critical technical and user interaction problems. For this, and keeping in mind the system was in early development stages, the experiment was conducted with blindfolded subjects performing the required actions.

The step-by-step of the experiment was the following:

1. A total of 15 subjects participated, 13 were registered in the database, leaving 2 to act as unknown;
2. For each registered user, 5 pictures were taken: 1 from a very short distance and 4 from the threshold distance. Of these 4, two were sideways (one for each side), one was frontal with a normal expression and one was frontal with a smile.
3. A participant was chosen to act as a blind user: first, they were taught how to open the GFR application, then they were blindfolded and, finally, they received a cane and instructions on what to do next.
4. In silence, four random participants were chosen to be placed in a short distance of the blindfolded persons. The only requirement was that one of these four was unknown in the database. They were positioned side-by-side, with their backs to a white wall (the same where the samples were taken).
5. Once the blindfolded subject was asked to start, the timer was set off and he/she had to enter the GFR application and recognize each of the four people in front of them, by their name or as unknown. To facilitate, the blindfolded user started facing the four people to be recognized and was positioned in the threshold distance from them.
6. For each person the blindfolded user recognized, s/he had to say aloud who s/he understood that person was. This was necessary so that the accuracy rate could be calculated in cases where framing issues caused different feedbacks to be given about the same person, for instance. Once all four people were recognized, the blindfolded user indicated they were done, and the timer was stopped.
7. The participant was kept blindfolded and taken back to the starting position. Steps 4 to 6 were repeated twice, with other two different groups of four people.
8. Steps 3 to 7 were repeated with a different blindfolded subject.

The previously described procedure was followed, except for the last blindfolded subject; the smartwatch’s battery ran out before the round with the last group could be completed. Additionally, another participant gave up before recognizing all four people, since s/he was not able to find one of them. Taking these two cases into account, in the end the experiment amounted to a total of 55 predictions. 46 of these were correct, giving an accuracy rate of 83.64%. Therefore, in terms of algorithms the GFR system presented a high accuracy rate and a satisfactory performance.

Regarding the user interaction, several problems were raised during the recognition stages, especially considering the context of accessibility. The main complaints revolved around the audio feedback, as it presented only two types of feedback: one to indicate the application was framing a person's face and another to provide the result of the recognition (the person's name or "unknown"). The "framing" feedback is a clue that the user needs to keep the wristwatch still, so that the system can analyze the captured face and, a few seconds later, provide the result of the recognition. However, the "framing" feedback was sometimes a false clue, either because the camera was not capturing a face or because the face being captured could not be analyzed. This caused frustration, as the blindfolded participant had to keep the arm elevated and bent at the elbow, to point the wristwatch's camera forward. Fatigue was another issue reported by all users that were blindfolded, since after each round it became more and more tiresome to keep the arm elevated.

Despite these problems, a positive aspect of the user interaction was found by analyzing the times the blindfolded users took in each round of recognizing a group of four people. As it is possible to see in Table 1, every participant had their worst performance in their first round, when they were still learning to use the GFR application. Then, most of them have their best performance on the second round and an average one on the last round. "Blind 3" was an exception because the application crashed on his last round, costing him some time. However, it is interesting to note that the average time for Round 2 was very close to the time of the specialist (researcher that was already well familiarized with the system and performed one round within the shown time). Additionally, the average for the first round is the highest and the average for the last round is the intermediate. Therefore, the decrease of average times from Round 1 to Round 2 indicates that the later interactions were easier, suggesting the system is easy to learn how to use. The increase in average times from Round 2 to Round 3 suggests the already mentioned fatigue issues.

Finally, the matter of the battery running out should be addressed. The experiment lasted about 2 hours, including the time taken to register the 13 users in the database. Considering that the GFR system is intended to serve as an assistive technology for the visually impaired, battery life is a critical issue. However, we highlight the fact that the smartwatch's screen was turned on the entire time, to allow the researchers to analyze the application's behavior. In a real contexts of use the screen would most likely be used very sparingly, increasing the time of battery life.

Table 1: Time taken for each round of people recognition.

	TIME (HH:MM:SS)		
	Round 1	Round 2	Round3
Specialist	00:01:29		
Blind 1	0:03:45	0:01:54	0:02:00
Blind 2	0:02:36	0:02:00	0:01:30
Blind 3	0:02:02	0:01:23	0:03:16
Blind 4	0:04:26	0:01:17	0:01:24
Blind 5	0:02:05	0:01:20	
Total	0:16:23	0:07:54	0:08:10
Average	0:03:17	0:01:35	0:02:02

6 CONCLUSIONS AND FUTURE WORK

In this paper we have described a real-time face recognition system built into a smartwatch with limited hardware and that features a 1.9 megapixel camera on its bracelet. The developed system detects the face captured by the camera and then performs the face recognition, emitting an audio feedback that identifies a recognized person or an unknown person. To run on the watch limited hardware, a variation of the K-NN algorithm was used for the face recognition approach. Finally, a pilot study was conducted to provide a preliminary evaluation of the GFR application. This evaluation included not only aspects of performance and user interaction, but also the design of the experiment itself, so that it is well refined when users with real disabilities are included in the studies.

In the pilot experiment, the system showed a satisfactory performance, with a high accuracy rate of 83.64%. The careful reader might have noticed that we used the K-NN recognition directly over the HOG features, which are on a high-dimensional space. This is quite unusual compared to what the literature describes, as the K-NN (or any other classifier) is usually applied after a dimensionality reduction stage, such as a PCA. The dimensionality reduction makes the system more robust, since everything is far from everything in a high-dimensional space. We avoided the PCA at this point because a PCA learns the subspace of interest from the training set. We are currently studying alternatives for a vanilla PCA, such as a self-updating PCA. This would use new exemplars, registered as the system performs, to estimate a more realistic subspace of operation. This will allow the system to start with a preregistered dataset, and improve its performance as it is used.

There is a lot of room to improve the actual accuracy of the system - we might be able to use more sophisticated face detection algorithms or classifiers, and even use techniques of hallucinating exemplars from the existing data, to make the system more robust to noise and illumination conditions. Nevertheless, we can strongly declare that our objective in this paper has been reached — it is technically possible to make a real-time robust face recognition system running exclusively on the low-performance hardware of the smartwatch.

Additionally, in terms of user interaction, the experiment was important to show usability and ergonomic issues that need to be addressed before people with actual visual impairments are involved. The feedback that indicates a face is being framed needs more work so that it becomes a more precise clue as to where the user needs to point the smartwatch's camera. This is important not only to allow the system to be used as an assistive technology, but also to alleviate the fatigue issue reported by the participants. Other potential place for future enhancement concerns the feedback interface to get data from people´s faces, which still must be made accessible for use by blind and low-vision people.

Finally, we propose challenges for future work, including wearable systems for objects recognition, textual information recognition (e.g. signs, symbols) and a gesture recognition like Porzi et al. (2013), but processed within the smartwatch itself. Furthermore, we will conduct experiments to better analyze the system's energy consumption. Also, experiments with visually impaired users will be used to further evaluate and improve the system as an assistive device.

ACKNOWLEDGEMENTS

The authors wish to express their gratitude to all the volunteers who participated in the experiments in this study, and also for Samsung Research that loaned the hardware equipment. LSBN receives a Ph.D. fellowship from CNPq (grant #141254/2014-9). VRMLM receives a Ph.D. fellowship from CAPES (grant #01-P-04554/2013). MCCB, ARR and SKG receives a Productivity Research Fellowship from CNPq (grants #308618/2014-9, #304352/2012-8 and #308882/2013-0, respectively). This work is part of a project that was approved by Unicamp Institutional Review Board CAAE 31818014.0.0000.5404.

REFERENCES

Ahonen, T., Hadid, A., and Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041.

Astler, D., Chau, H., Hsu, K., Hua, A., Kannan, A., Lei, L. Nathanson, M., Paryavi, E., Rosen, M., Unno, H., Wang, C., Zaidi, K., Zhang, X., and Tang, C. (2011). Increased accessibility to nonverbal communication through facial and expression recognition technologies for blind/visually impaired subjects. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 259–260.

Belhumeur, P., Hespanha, J., and Kriegman, D. (1997). Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720.

Chen, X., Flynn, P., and Bowyer, K. (2003). PCA-based face recognition in infrared imagery: baseline and comparative studies. In *Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, pages 127–134.

Cover, T., and Hart, P.: Nearest neighbor pattern classification. (1967). *IEEE Transactions on Information Theory*, 13(1):21–27.

Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 886–893.

FreevoxTouch (2014). The only smart watch in the world for the visually impaired. Available: http://myfreevox.com/en/.

Fusco, G., Noceti, N., and Odone, F. (2012). Combining retrieval and classification for real-time face recognition. In *2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance*, pages 276–281.

Gordon, G. (1991). Face recognition based on depth maps and surface curvature. In *SPIE1570, Geometric methods in Computer Vision*, pages 234–247.

Hadid, A., Pietikainen, M., and Li, S. (2007). Learning personal specific facial dynamics for face recognition from videos. In *Analysis and Modeling of Faces and Gestures: Lecture Notes in Computer Science 4778*, pages 1–15.

Kistler, D. and Wightman, F. (1992). A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *Journal of the Acoustical Society of America*, 91(3):1637–1647.

Kramer, K., Hedin, D., and Rolkosky, D. (2010). Smartphone based face recognition tool for the blind. In *32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4538–4541.

Krishna, S., Little, G., Black, J., and Panchanathan, S.

(2005). A wearable face recognition system for individuals with visual impairments. In *The Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 106–113.

Li, B., Mian, A., L., W., and Krishna, A. (2013). Using kinect for face recognition under varying poses, expressions, illumination and disguise. In *IEEE Workshop on Applications of Computer Vision*, pages 186–192.

Luxand, Inc. (2013). Detect and Recognize Faces with Luxand FaceSDK. Updated on August 27, 2013. Available: https://www.luxand.com/facesdk/

Manduchi, R. and Coughlan, J. (2012). (Computer) vision without sight. In *Communications of the ACM*, 55(1):96–104.

NEUROtechnology. (2014). VeriLook SDK: face identification for stand-alone or web applications. Updated on April 15, 2014. Available: http://www.neurotechnology.com/verilook.html.

Organization, W. H. (2013). Visual impairment and blindness: Fact sheet n.282. Available: http://www.who.int/mediacentre/factsheets/fs282/en/

Porzi, L., Messelodi, S., Modena, C. M., and Ricci, E. (2013). A smart watch-based gesture recognition system for assisting people with visual impairments. In *Proceedings of the 3rd ACM International Workshop on Interactive Multimedia on Mobile*, pages 19–24.

Pun, T., Roth, P., Bologna, G., Moustakas, K., and Tzovaras, D. (2007). Image and video processing for visually handicapped people. *EURASIP Journal on Image and Video Processing*, 2007:025214(5):4:1–4:12.

Schwartz, W., Guo, H., Choi, J., and Davis, L. (2012). Face identification using large feature sets. *IEEE Transactions on Image Processing*, 21(4):2245–2255.

Tanveer, M., Anam, A., Rahman, A., Ghosh, S., and Yeasin, M. (2012). Feps: A sensory substitution system for the blind to perceive facial expressions. In *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 207–208.

Tistarelli, M. and Grosso, E. (2010). Human face analysis: From identity to emotion and intention recognition. In *Ethics and Policy of Biometrics: Lectures Notes in Computer Science 6005*, pages 76–88.

Turk, M. and Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86.

Viola, P. and Jones, M. (2004). Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154.

Watanabe, H., Terada, T., and Tsukamoto, M. (2014). A sound-based lifelog system using ultrasound. In *Proceedings of the 5th Augmented Human International Conference*, 59:1–59:2.

Wilder, J., Phillips, P. J., Jiang, C. and Wiener, S. (1996). Comparison of visible and infra-red imagery for face recognition. In *Proceedings of the 2nd International Conference on Automatic Face and Gesture Recognition*, pages 182–187.

Wiscott, L., Fellous, J., and Malsburg, C. (1997). Face recognition by elastic buncg graph matching. In *IEEE Transactions on pattern analysis and machine intelligence*, pages 775–779.

Zhao, G. and Pietikainen, M. (2007). Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):915–928.

Zhao, W., Chellappa, R., Phillips, P. J., and Rosenfeld, A. (2003). Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399–458.