A Spectral Information Coding System for the Visually Impaired

Shao-Ting Chiu[1,2], Yi-Hsuan Wu[1]

[1]Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan

[2]Department of Dentistry, National Yang-Ming University, Taipei, Taiwan

Correspondence: b03901045@ntu.edu.tw | +886937857127

*This paper adheres to the panel- “Smart services of smart city”.

ABSTRACT

The visually impaired often have trouble in urban navigation or having eye contact in social events due to the disability of target tracking. In this study, we focused on smart service for the visually-impaired and designed a wearable navigation tool which contains object recognition, dorsal tactile display and eye-controlled input. The image captured with the frontal camera is processed by a neural network with two hidden layers, and the spatial information of targets is displayed with a wearable vibrator motor array. Also, in order to increase the processing speed, frontal image is narrowed down to the area around the eye focus as input. Our prototype and testing suggest that users can locate the selected targets via tactile display with this system.

INTRODUCTION

The visually impaired usually navigate the surroundings by sound or with tactile sense (e.g. by touch or using a cane). Thus, it is difficult to locate signs such as pedestrian crossing and bus stop at a distance, which are essential for safety and navigation in the city. On the other hand, the visually impaired often fail to have eye contact with people, which is a disadvantage during social events. From the interviews we had with the visually impaired living in Taipei City, we concluded that there is a strong demand for the visually impaired to locate certain objects that are essential for urban navigation or social events.

Smart glasses such as Pivothead can recognize and interpret images by artificial intelligence, enabling the visually impaired to understand who and what is going on around them[1], [2]. However, locating the specific objects in real time is demanding for this device because they convey graphical information solely by vocal speech.

Tactile sense provides another option for the visually impaired to understand the environment around them in real time. In human skin, there are four channels of mechanic receptive afferents thought to form four distinct sensory of touch[3], [4]. Though the way these tactile receptors encode mechanical vibration is still under discussion[5], [6], a study done with vibration frequency of 10 Hz suggests that the tactile response to vibrational amplitude is typically a piece-wise linear function[7]. Thus, to convey the information via tactile sensation, the vibrational amplitude should work on the linear regions.

Previous works employing vibrating vest usually contains an array of vibration motors, multiple sensors include lidar, ultrasound and infrared[4], [8], [9]. The aim of these designs is to enable the visually impaired to be aware of obstacles such as an approaching car and can take corresponding actions in advance. Although these designs are able to sense the environmental condition via multiple sensors, the original data is incomprehensible without filtering by artificial intelligence. Moreover, due to the noise of vibrational display and the uncertainty of tactile response, the information capacity of these systems may not be enough to correctly convey environmental information[4], [10], [11].

To sum up, environmental detection, tactile display, and artificial intelligence are essential for the transmission of spectral information to the visually impaired. We thus proposed a wearable spectral coding system to locate environmental objects in real time via a vibration motor array.

DESIGN AND MATERIALS

1. *The Proposed System*

Figure 1 depicts the workflow of the proposed system. Calculation 1 (Cal 1) describes the status of the user's eye (Fig. 1): if the eyes are closed, which means the user is not aiming at any objects, the vibration motors will be switched off; on the other hand, if the user's eyes are opened, the frontal image will be captured and cropped around the area at which the eyes aim. Calculation 2 (Cal 2) then locates the regions of the objects which the user has previously selected by pouch panel or microphone (Fig. 1).

Finally, the spectral information of the targets are encoded into a vibration pattern. To indicate the regions of targets, the duty cycle of each motor is determined by Eqn. 1,

$$\text{Duty cycle} = \frac{Area\ of\ the\ Object\ \ within\ \ a\ \ Square}{Total\ Area\ of\ a\ \ Square} \times\ 100\%, \qquad (1)$$

where an input image is divided into several squares corresponding to the vibration motor array.

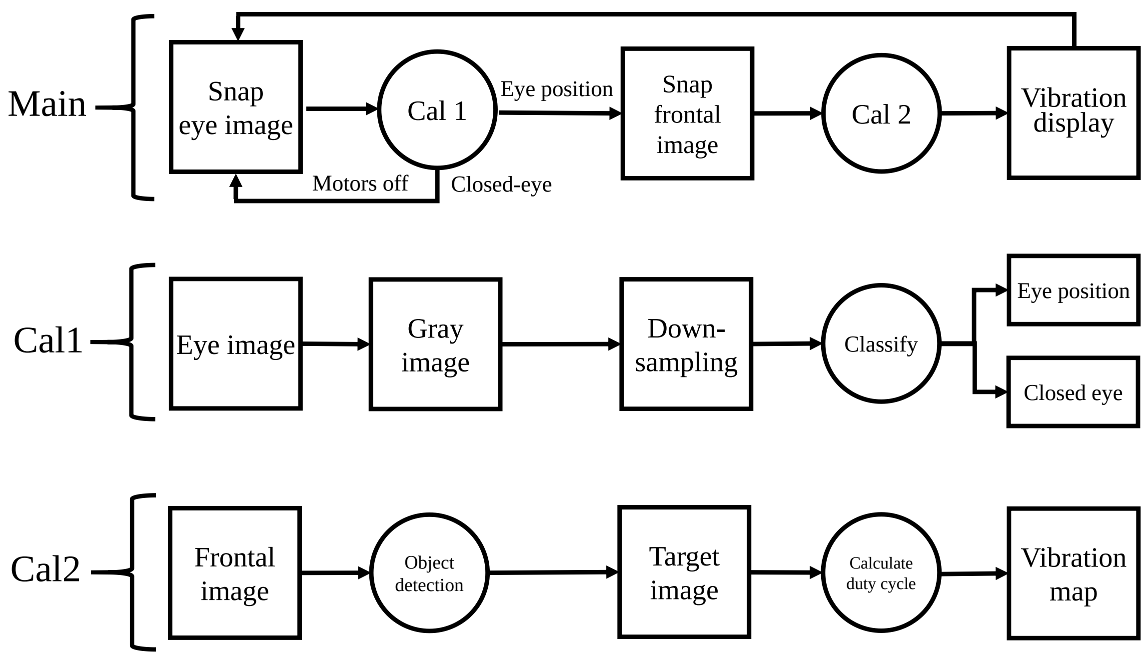

Fig. 1. Block diagram of the proposed spectral coding system. Calculation 1 (Cal1) determines whether eyes are closed or are at certain positions. Calculation 2 (Cal2) locates the regions of targets in the frontal image, and encodes the spectral information into a vibration map, a matrix used to record duty cycles.

2. *Materials*

The prototype is based on Raspberry Pi 3 Model B, a small single-board computer. Two USB webcams worn on the head are connected to the Raspberry Pi: the frontal camera captures the image in front of the user, while the rear camera captures the image of a user's eyeball. Two images both determine a vibration pattern. In addition, we implemented a vest with a 5×5 vibration motor array attached to the user's back (Fig. 2). All the programs are written in Python 2.7.11.

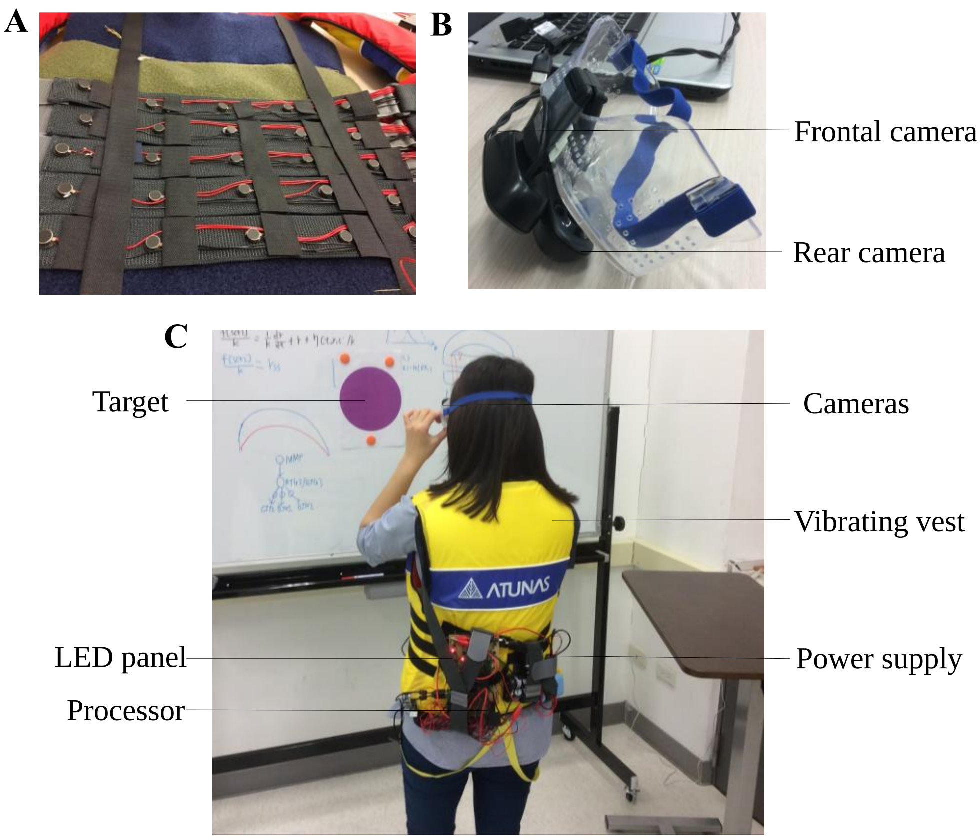

Fig. 2. (A) The 5×5 vibration motor array (B) Smart glasses with two cameras. Frontal camera captures the frontal image, and rear camera captures the image of user's left eye. (C) A tester wore in the prototype. The system was set to track a purple round and displayed vibration on the back (see SOM movie 1. and 2.).

RESULTS

1. *Closed Eye Detection with Artificial Neural Network*

The closed eye detection is to provide part function of hands-free control in the proposed system (Fig. 1).

Due to the limited speed of CPU on Raspberry Pi 3 Model B, the typical eye tracking algorithm, Haar-cascade[12], can't accomplish ideal real-time detection (0.25 frame per second on our platform). We thus constructed an artificial neural network with two hidden layers for closed eye detection at the processing speed of 10 fps (frame per second). The algorithm works with test accuracy of 100% and train accuracy of 94.3% (Fig. 3).

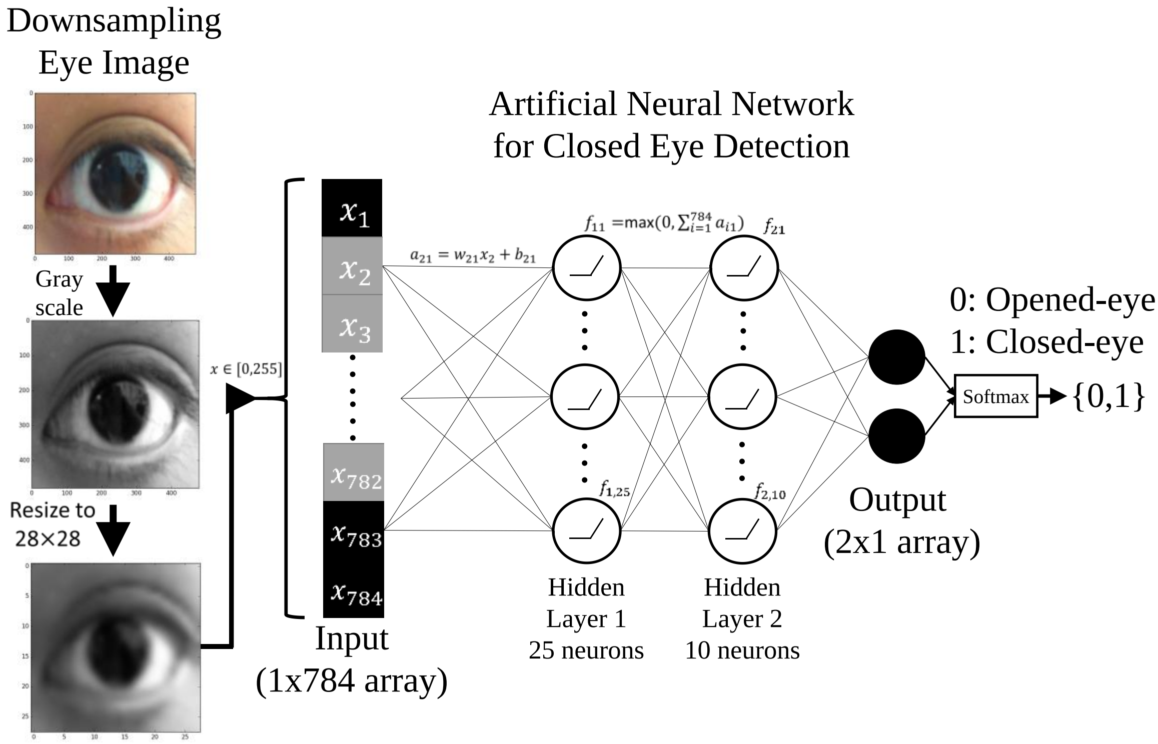

Fig. 3. Closed eye detection. An image of eye ball is classified by artificial intelligence. First, the eye image is transformed to gray scale and resized to speed up the process. Second, 784 pixels are featured by an artificial neural network with two hidden layers and RELU neurons[14]. Finally, this algorithm can work with an accuracy of 94.3% and at 10 fps (frame per second) in Raspberry Pi 3 Model B. x_i denotes the ith pixel intensity.

2. Real-Time Object Tracking

We implemented the prototype of real-time object tracking algorithm (Fig. 4A and Fig. 1). Object tracking takes place only with the user's eyes are opened. First, frontal image is taken as input, and our algorithm filters out other colors which don't belong to the target. Second, the regions of targets are recognized by an artificial neural network with structure similar to closed eye detection (Fig. 3). Third, the duty cycle of each vibration motor is calculated according to the Eqn. (1). Finally, the vibration map stores the values of duty cycles in the following 5×5 martix.

$$VibrationMap = \begin{bmatrix} D_{1,1} & \cdots & D_{1,5} \\ \vdots & \ddots & \vdots \\ D_{5,1} & \cdots & D_{5,5} \end{bmatrix}, \qquad (2)$$

where $1 \geq \mathrm{D} \geq 0$, $D_{i,j}$ denotes the working duty cycle of the vibrator on row i and column j (Fig. 2A).

In this prototype, we set the system tracking the purple symbol of zero, and tested the performance under several conditions (see SOM movie 3). We found that the system can still recognize the target even if there are objects with either

A

B

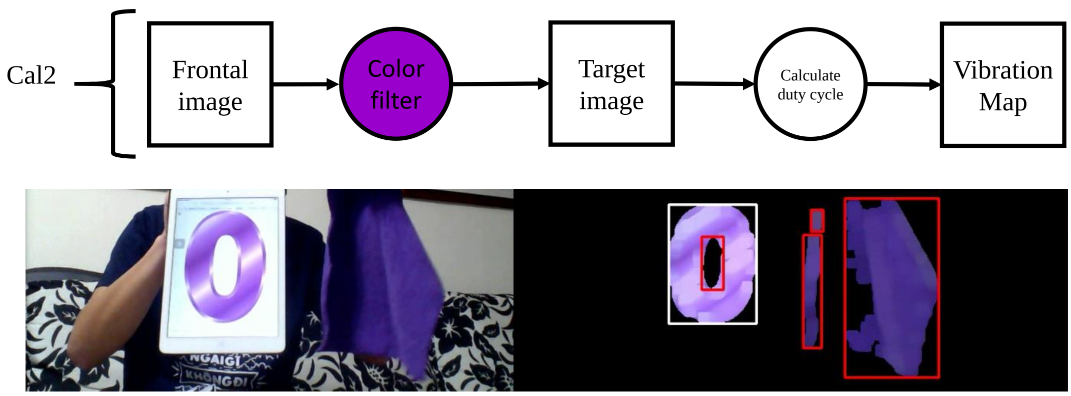

Fig. 4. Real time object tracking. (A) The block diagram of calculation 2 (Cal 2, Fig. 1) (B) The left screen displays an input image, and the right screen displays the processed one. First, the purple symbol of zero is selected as target. Second, the software filters out other colors in the input image. Then, an artificial neural network recognizes the symbol. Finally, the recognized target is circled in white and the rejected candidates are circled in red. (See also SOM Movie 3.)

similar color or similar morphology. This system can thus precisely recognize the regions of targets.

Furthermore, the processing speed is also a crucial parameter of performance because objects in a city like cars often move at high speed. Our previous Harr-cascade algorithm[12], which works at 0.3 fps (frame per second) on our platform, fails to reach the requirement. Fortunately, the later algorithm works in 50 fps on Raspberry Pi 3 Model B, which is quick enough for the real-time detection.

3. The 5×5 vibration motor array

The motor array vibrates at the frequency of 40 Hz[7], and the duty cycle is set according to the vibration map (Eqn. 2 and Fig. 4A). We further tested the performance under the condition of moving target and stationary target by a tester wearing the device, and observed the vibration pattern via a 5×5 LED panel connected to motors (Fig. 2C. See also SOM movie 1. and 2.). Results suggest that the user can locate the direction of the target and estimate the distance based on the known target area.

Also, the duty cycle of motor vibration corresponds with the outline of the target. When the target is at a distance, the corresponding duty cycle enables user to estimate the distance.

To sum up, this system has the potential for enabling the visually impaired to locate objects via tactile sense.

DISCUSSION

To correctly convey spectral information of certain object, we proposed a navigation vest which includes a programmable vibration motor array, microprocessor, a pair of cameras for capturing frontal image and user's eyeballs. Also, an artificial neural network inside the system is used to recognize the target, and the other neural network recognizes the status of user eyeballs to decide the input region of frontal image. The proposed eye-controlled design enables users to choose the input region in will and to hold eye contact with the target;

furthermore, we constructed the prototype including closed eye detection system, object tracking, and a wearable vibrating vest.

With the rapid development of artificial intelligence, the deep convolutional neural network can achieve reasonable performance on hard visual recognition tasks or even exceeding human performance in some domains[13]. This technology provides potential for large-scale image recognition which is essential for the navigation of the visually impaired.

Moreover, speech recognition enables user to select the desired object by just speaking a term. Implementing this technology will make our proposed system smarter.

Existing products such as Google Glass[15] and Pivothead[1] employ embedded cameras, microphone, voice recorder and small processer, which provide possible solutions to make our system lighter and simpler to be used as mobile navigation tool.

Therefore, with the aid of deep convolutional network, speech recognition and smart glasses, the proposed system will be smarter and more user-friendly for the visually impaired living in cities.

CONCLUSION

By testing the prototype theoretically and practically, we conclude that our proposed spectral coding system enables users to locate selected targets without the need of eyesight, and includes hands-free control. This system can thus provide the visually impaired with the spectral information in front of them via tactile display. This information can be further used to navigate in cities and hold eye contact with people during social events. Thus, in this study, we propose a design of smart service for the visually impaired.

SUPPORTING ONLINE MATERIAL

http://dx.doi.org/10.17632/p6hxtmdpsp.2

Movie 1. Vibrating vest- detecting moving target

Movie 2. Vibrating vest- detecting stationary target

Movie 3. Real-time object tracking

REFERENCES AND NOTES

[1] "SMART — Pivothead Wearable Imaging." [Online]. Available: http://www.pivothead.com/smart-2/.

[2] Sujith B and Safeeda V, "Computer Vision-Based Aid for the Visually Impaired Persons-A Survey And Proposing New Framework," *Int. J. Innov. Res. Comput. Commun. Eng. (An ISO Certif. Organ.*, vol. 32972, no. 5, 2007).

[3] S. J. Bolanowski, G. A. Gescheider, and C. M.Checkosky, "Four channels mediate the mechanical aspects of touch.", *J Acoust Soc Am*, vol. 84, pp.1680-1694, 1988. [PubMed: 3209773]

[4] B. W., L. S.Paul Bach-Y-Rita, Carter C. Collins, and Frank A. Saunders, "Vision Substitution by Tactile Image Projection," *Nature*, vol. 221, pp. 963–964, Nov. 1969.

[5] G. G. Lin and J. G. Scott, "Tactile intensity and population codes," *Behav Brain Res.*, vol. 100, no. 2, pp. 130–134, 2012.

[6] Y. Vázquez, E. Salinas, and R. Romo, "Transformation of the neural code for tactile detection from thalamus to cortex.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 110, no. 28, pp. E2635-44, Jul. 2013.

[7] K. Johnson, "Reconstruction of Population Response to a Vibratory Stimulus in Quickly Adapting Mechanoreceptive Afferent Fiber Population Innervating Glabrous Skin of the Monkey." *J Neurophysiol*, vol. 37, pp.48-72, 1974 [PubMed: 4204567]

[8] "Tactile Navigation Tools | Next Generation Assistive Technology." [Online]. Available: http://tactilenavigationtools.com/.

[9] D. Tudor, L. Dobrescu, and D. Dobrescu, "An Ultrasonic Electronic System for Blind People Navigation," pp. 5–8, 2015.

[10] M. Hollins and S. J. Bensmaïa, "The coding of roughness.," *Can. J. Exp. Psychol.*, vol. 61, no. 3, pp. 184–195, Sep. 2007.

[11] R. Ackerley, I. Carlsson, H. Wester, H. Olausson, and H. Backlund Wasling, “Touch perceptions across skin sites: differences between sensitivity, direction discrimination and pleasantness.,” *Front. Behav. Neurosci.*, vol. 8, p. 54, Feb. 2014.

[12] S.Soo, “Object detection using Haar-cascade Classifier.”

[13] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C.Berg, and L. Fei-Fei, “ImageNet Large Scale Visual Recognition Challenge,” *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec.2015.

[14] M.Abadi et al., TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL http://tensorflow.org/. Software available from tensorflow.org.

[15] Shimpali Deshpande, M. Geeta, and Chaudhari. Google Glass. *Int. J. Sci. Eng. Res.* 4**,** (2013).

[16] **Acknowledgments:** We are grateful to Wei Fang and Chun-Lu Tsai for insightful business model, interviews with the visually impaired and discussions; Ting-Yu Wu for providing legal advisory; Dr. An-Chi Wei for providing test area. This project was supported by Taidah Entrepreneurship Center of National Taiwan University, and also named “Visual Force Project” in NTU Challenge Contest. The product information can be accessed at NTU Challenge official website: https://ntuchallenge.com/teams/show20.html?team_id=hc_09#header