2011 International Conference on Computer Communication and Management
Proc .of CSIT vol.5 (2011) © (2011) IACSIT Press, Singapore

Signage Recognition Framework for Visually Impaired People

Oi-Mean Foong and Nurul Safwanah Bt Mohd Razali
Universiti Teknologi PETRONAS
Bandar Seri Iskandar, Tronoh 31750, Perak, Malaysia

Abstract. This paper presents a signage recognition framework for Malaysian Visually Impaired People (VIP). Independent navigation is always a challenge to visually impaired people whenever they attempt to find their way in a building (i.e. an office, a seminar room) and other amenities (i.e. an exit, ladies/gents toilets) even with a walking stick or a guide dog. The proposed framework will capture an image of a public signage and transform it into a text file using Otsu's optical character recognition method. The text file will be read by a speech synthesizer that tells the visually impaired people what the image is. This framework does not require huge database of the signage but only the character database. The framework is being tested on 5 blind folded subjects to check on its performances. Experimental results yield a recognition rate of 86.0%. This proof of concept shows that signage recognition could be easily achieved using optical character recognition and voice synthesizer.

Keywords: Optical character recognition, public signage, visually impaired people, way finding.

1. Introduction

Visually impaired is a state where the partially blind people have difficulties on seeing things which in some degree can be considered as handicap. People with normal long-sighted and myopic are not considered as handicap or visually impaired [1]. There are an estimated 45 million blind people and 135 million visually impaired people worldwide [6]. On the statistics by gender, female has the significantly higher risk to be visually impaired than male, in every parts of the world and also every age group. There are many factors contributing to visually impairment. For examples accident, glaucoma, diabetes, coloboma and the most common cause is age-related macular degeneration [2]. Visually impaired use different ways and methods to make their life easier to move around doing chores, learn and many more [3]. The most common way used by a visually impaired or a blind person to navigate around is by using a walking stick. There are others who rely on guide dog to bring them to places that they want to go. The guide dogs have been specially trained to help the blind or the visually impaired to move around and guide them but this dog can only go to places that they already being trained.

Normally, blind people had to acquire help from other people. They cannot be fully independent due to vision weaknesses. They had to depend on their caretakers or children to help them to go to places, etc. This problem has posed a major barrier for them to live independently.

The objectives of this paper are:

- To propose a framework of optical character recognition (OCR) on public signage for VIP.
- To integrate the public signage recognition with voice synthesizer.

The proposed optical character signage recognition framework is capable of detecting and recognizing signage which consists of alphanumeric characters, i.e. all upper or lower case alphabets from A to Z and digits from 0 to 9. The framework will assist in designing better signage recognition prototype system which helps the visually impaired people to read the public facilities signage like toilet, mosque, etc. As a result, it could reduce the dependency of the VIPs on their caretakers.

This paper is organized as follows: Section 2 contains the motivation behind the research. Section 3 describes the existing computer vision-based way finding for the VIPs or blind persons. Section 4 presents the details of the proposed signage recognition framework. Section 5 includes the experiments carried out and the results. Last but not least, section 6 concludes the findings of the research work.

2. Research Motivation

This research project was initiated after a site visit to National Council of the Blind Malaysia (NCBM) in Kuala Lumpur. An interview was conducted on how the visually impaired people (VIP) use computers in general. From the survey, it is learned that the blind or VIP use assistive technology such as courseware and/or gadgets in their daily lives [4], [5]. One of the outcomes of the preliminary survey is that the VIPs desperately need affordable gadgets in detecting and recognizing public signage in assisting them to navigate around various places. This is especially pertinent to VIPs who go to work just like sighted professionals and they enjoy absolute independence without the help of caretakers.

3. Existing Systems

Ezaki Nabuo et al. used the PDA, charged coupled device (CCD) - camera, and also the voice synthesizer on building this system [7]. The camera that they use had to be on the user's shoulder. The signage will be randomly taken by the camera but may not be the signage that the users want.

H. Simon et al. presented a framework for travel by visually impaired people in order to provide a basis on which to develop future mobility and orientation systems to support such travelers [8]. The 'travel task' is reviewed, and then analyzed in the context of visual impairment, leading to the development of a simple framework as a means of examining travel and travel aids for visually impaired people so that inadequacies in the provision of these travel aids can be addressed. Similar idea is adopted to help the VIP to travel around the city using mobile phone [9].

S. Michael et al. designed an assistive tool to help the visually impaired by using hand held optical scanner that can be used for manual scanning to printed text with standardize font that using the character recognition to read the printed materials and the output will be the voice synthesizer [10]. But the gadget failed to ensure straight line motion as the reader moves the scanner either backward or forward as the user scanned across the document.

The vOICe acts as a research vehicle for the cognitive sciences to learn more about the dynamics of large-scale adaptive processes, including attention and expectation guided bottom-up and top-down learning processes [11]. From neuroscience research it was shown that the visual cortex of even adult blind people can become responsive to sound and sound-induced illusory flashes can be evoked in most sighted people. The vOICe is built on live video that comes with head-mounted camera that is encoded with sound. The gadget is quite expensive even for middle class people to buy it. Besides, the device is just to recognize the objects without helping the blind people to read the text in the object.

The *Cell Phone-Based Way Finding* project was pioneered by the researchers in the University of California. The proposed system was tested using Nokia 7610 by the real blind person and it would tell the blind users what was the destination. The camera in the cell phone could detect the signs even though it was few meters away from the signage [12] [13] [14].

The signage was specifically designed in such a way that that the cell phone camera can read the signage which had barcode on it. The prototype searched the database of barcodes in the phone and compared it with the detected barcode that was displayed near the signage. This project is very useful for the visually impaired as it is able to read the signage few meters away since blind people cannot really see where the barcodes is located. However, the system can only read the barcode that is installed near the signage. At times if the barcode is missing from the signage, the visually impaired people might not be able to detect the signage and use the facilities

4. Proposed Signage Recognition Framework

The proposed signage recognition framework consists of digital camera as shown in Fig. 1. Matlab image processing toolbox and voice synthesizer using SAPI (Speech Application Program Interface) are deployed. This framework is divided into two parts: (1) Image to text conversion where the optical character recognition algorithm is used and (2) From text to voice conversion where it will tell the visually impaired or the blind person what is written on the signage.

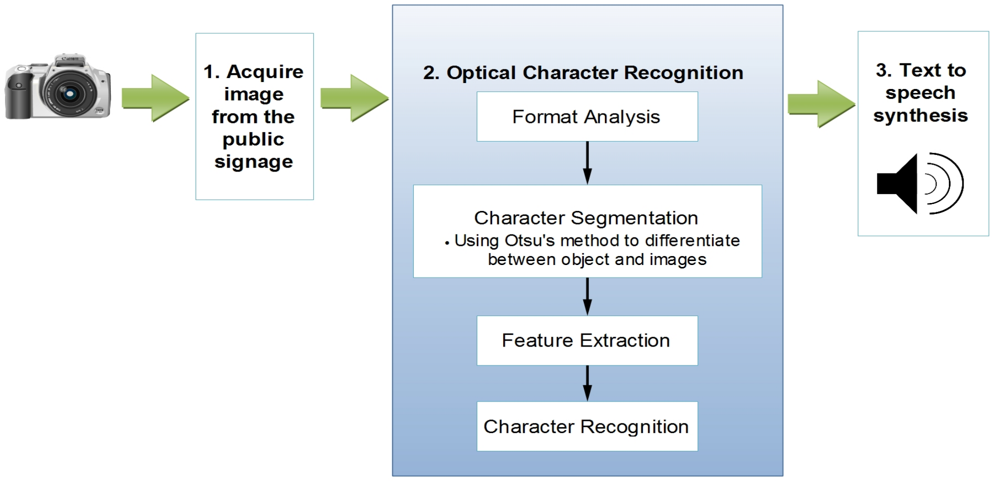


Fig. 1: Proposed Signage Recognition Framework

4.1 Image Acquisition

Image acquired will be a signage which contains either alphabets like ladies, toilet, school or any alphanumeric characters. The user captures the signage using camera and save it in .jpg file. After the image has been captured by the camera, the image will be processed and transformed into a text file. The recognition process of character images is shown in Fig. 1.

4.2 Format Analysis

The image taken will be analyzed to find the appropriate character image. For our proposed system, the image taken will go through the process of adjusting the image contrast and intensity from colored image to make the image clearer using the following code in MatLab:

imagen2 = imadjust(imagen, [0.3 0.5], [ ]);

The image will be transformed to grayscale image.

4.3 Segmentation

In segmentation process of the character which is based on the R.G. Casey [15] research, this step requires to answer "What constitutes a character?" It is the process to determine the element of an image. In segmenting image and converting it to the binary image, the Otsu's method is applied to help in differentiating the foreground and background images [16]. Eq. 1 shows the Otsu's formula to find the foreground and background variances:

$$\delta_w^2 = w_b\,\delta_b^2 + w_f\,\delta_f^2 \qquad (1)$$

where

δ_w^2 = Total variance for background and foreground.

w_b = Weight for background. Values found by adding the total background threshold value divide by the total pixel.

δ_b^2 = Variance for background. Values found by multiple each background weight value with means and divide it with total threshold value.

w_f = Weight for foreground. Values found by adding the total foreground threshold value divide by the total pixel.

δ_f^2 = Variance for foreground. Values found by multiply each foreground weight value with means and divide it with total threshold value.

4.4 Feature Extraction

The feature extraction process is applied to differentiate between the object and its background and convert it to binary image. The binary image will be filtered by removing noise and also capture the essential features of the image. The image which has less than 30 pixels will be eliminated and filtered as noise. The binary image will be represented as a matrix.

4.5 Character Recognition

The binary image created will be compared and matched with the template file that has been created. An attempt will be made to find the nearest character template which is similar to it. The image that has earlier been recognized will be saved in the text document as text.txt.

4.6 Text to Voice

The text file that has been created will be fed into Microsoft Speech Application Program Interface (SAPI) to tell the user what the signage is as voice output.

5. Results and Discussion

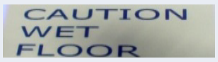
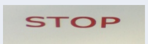
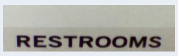
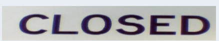
5.1 Experimental Setup

In this research work, public signage which consists of English alphabets is considered. Images of common public signage such as caution wet floor, bus stop, danger, closed, open, exit, taxi, stop, restrooms and road signage were taken using built-in camera from a mobile phone. These images were stored in a database in a computer for comparing with the input images randomly captured by blind-folded subjects. A total of five normal sighted subjects (3 females and 2 males) had volunteered in the experiments.

5.2 Findings and Discussion

Each blind-folded subject captured an indoor signage using digital camera and save in .jpg file format. The image was transmitted to a computer for comparison purpose. Table 1 shows the results of signage which are correctly recognized for the proposed framework.

Table 1 System Performance Evaluation Results

Signage	Signage correctly recognized
	Yes
	Yes
	Yes
	Yes
	No
	Yes
	Yes
	Yes
	Yes
	Yes

The public signage recognition is evaluated using Eq. 2:

$$True\ \ Positive\ \ Rate\ (TPR) = \left(\frac{No.of\ correct\ \ signage}{Total\ \ no.of\ \ signage} \right) X\,100\% \qquad (2)$$

Table 2 shows the percentage of True Positive Recognition Rate for each blindfolded subject. Based on observation, each image taken must be in capital letters or can be a bit slathered but not slant more than 30 degrees. Otherwise too much lighting by using flash will cause the image to be blurred or defective. The results reveal that the signage recognition framework attains an average recognition rate of 86.0%.

Table 2 Experimental Results on the Respondents

User	No. of Samples	Recognized Samples	Recognition Rate %
1	10	9	90%
2	10	10	100%
3	10	8	80%
4	10	9	90%
5	10	7	70%
Total	**50**	**43**	**86% (Average)**

6. CONCLUSION

The proposed framework with OCR for signage recognition is integrated into two parts, image to text process using the Otsu's method to differentiate the background and foreground object, and text to speech process using the SAPI. This work shows the proof of concept of our proposed computer vision-based way finding framework for assisting visually impaired people to independently access unfamiliar places. The framework has successfully used voice synthesizer to tell the VIP what the signage is when they are in doubt. For future improvement, this framework should be able to differentiate between the alphabet symbols and the text image correctly, and it is capable of reading any public signage in lowercase/upper cases even if it is slathered. Their dependency to ask people for help will be less as they could use assistive technology tool for way finding.

7. Acknowledgment

Special thanks to Moses Choo from the National Council of the Blind Malaysia (NCBM) for providing support in this research. We also thank all staff who either directly or indirectly involved in the project.

8. References

[1] Finish Federation of the Visually Impaired. Finland: Finnish Federation of the Visually Impaired (FFVI). http://www.nkl.fi/yleista/english.htm

[2] NHBVI, New Hampshire: Common causes of visual impairment. http://www.nhbvi.com/Agenda/Goal1/Book1/commoncauses.html

[3] American Council of The Blind. America : What is Braille? http://www.acb.org/resources/braille.html

[4] Suziah Sulaiman, Siti Nur Syazana Mat Saei, Oi Mean Foong, and Halabi Hasbullah, Understanding Domain Expert's Perspectives and Expectations in Assistive Technology, In: IEEE International Symposium on Information Technology (ITSIM 2010), IEEE Xplore, pp. 1164-1167, Jun 2010. oi:1109/ITSIM.2010.5561491.

[5] Mohahd Ghazali Tocheh. Siapa Bersama Mereka. National Council For The Blind Malaysia (NCBM). Laser Press Sdn Bhd, 2006.

[6] World Health Organization WHO: Visual Impairment and Blindness. http://www.who.int/mediacentre/factsheets/fs282/en/

[7] N. Ezaki, M. Bulacu, and L. Schomaker. Text Detection from Natural Scene Images:Towards a System for Visually Impaired Persons, In: 17th International Conference on Pattern Recognition (ICPR), IEEE Computer Society, pp.683-686, 2004. DOI: 10.1109/ICPR.2004.1334351.

[8] H.Simon , G.Peter , A Travel Flow and Mobility Framework for Visually Impaired Travellers,1990.

[9] C.James, M.Roberto, and S.Huiying, Cell Phone-based Way Finding for the Visually Impaired, 1st. International Workshop on Mobile Vision, pp.1-15, 2006.

[10] S.Michael, Y,Kipman, Hand Held Optical Scanner for Omni-Font Character Recognition , pp. 163-164, 1987.

[11] B.L. Meijer. The vOICe Next-Generation Human Vision Technology. http://www.seeingwithsound.com/about.htm

[12] C.James, M.Roberto, S.Huiying. Cell Phone-based Wayfinding for the Visually Impaired, 2005.

[13] B.S. Tjan, P.J. Beckmann, R. Roy, N. Giudice, and G.E. Legge, Digital Sign System for Indoor Way Finding for the Visually Impaired, 2005.

[14] Bae, K.S.; Kim, K.K.; Chung, Y.G.; Yu, W.P. Character recognition system for cellular phone with camera, In:29th Annual International Computer Software and Applications Conference, COMPSAC 2005, vol.1, n.o.2, pp. 539-544, July 2005.

[15] G.C. Richard. A Survey of Methods and Strategies in Character Segmentation, pp. 690-706, 1996.

[16] N. Otsu, A Threshold Selection Method from Gray Level Histograms, IEEE Transactions in Systems Man and Cybernetics, vol.9, pp.62-66, 1975.