SUGISAKA MASANORI

Journal of Robotics, Networking and Artificial Life
Vol. 9(1); June (2022), pp. 25–30
SSN (Online): 2352-6386; ISSN (Print): 2405-9021
https://alife-robotics.org/jrnal.html

Research Article

A Method of Navigating a Visually Impaired Person Using MY VISION

Joo Kooi Tan, Kohei Kitagawa
Department of Mechanical and Control Engineering, Kyushu Institute of Technology, Sensui-cho Tobata-ku, Kitakyushu, 804-8550, Japan

ARTICLE INFO

Article History

Received 24 November 2020
Accepted 02 November 2021

Keywords

MY VISION
Visually impaired
GBS, GPS, HOG
Random Forest

ABSTRACT

This paper describes a method of navigating a visually impaired person who walks on a sidewalk and goes through a pedestrian crossing by the use of an ego-camera system named MY VISION. The proposed method finds a walkable region on the sidewalk. Thus, at a pedestrian crossing, it finds a traffic light and judges an appropriate timing to cross the road by analyzing the color and flashing status of the signal. The walkable region is found employing Graph- Based Segmentation (GBS), and once a pedestrian crossing is reached through the use of a GPS signal, a traffic light is detected by use of the HOG feature and Random Forest. The effectiveness of the proposed method was verified by experiments. This implies usefulness of a MY VISION system

1. Introduction

According to the 2011 Ministry of Health, Labor and Welfare report [1], there are approximately 310,000 people in Japan who have a visual impairment. Looking at the world, it is predicted that the number of visually impaired people worldwide will triple within the next 40 years [2].

Accidents of visually impaired people are not uncommon, and there is a survey result that about 42% of the subjects experienced accidents with injuries while walking outdoors within the past 5 years. The accidents are caused mostly by automobiles [3]. It is easy to imagine that those accidents with cars may often happen at an intersection, which is one of the most dangerous places for a visually impaired person to cross.

A visually impaired person normally finds an intersection by Braille blocks and/or acoustic signs and judges the timing of crossing it acoustically. However, not all intersections are equipped with Braille blocks or acoustic instruments. Therefore, it is necessary to develop a technique for recognizing and analyzing traffic lights, for a visually impaired person to cross the intersection safely.

In order to detect traffic lights, a method using template matching and a method utilizing position information by Global Positioning System(GPS) or a map have been proposed [4-6]. However, all these studies are limited to the detection of traffic lights, and they do not analyze the possible timing for a visually impaired person to cross. It is vital for a visually impaired to receive the information about the timing as to when to cross. The information should be derived through the analysis of traffic light images, once it is detected.

Corresponding author's E-mail: etheltan@cntl.kyutech.ac.jp

In this paper, we propose a method of detecting a pedestrian traffic light at a pedestrian crossing by MY VISION and analyzing the states of the signal. MY VISION is an ego-camera system employing a body/head mounted camera. It can be a virtual eye of a visually impaired person or the third eye of a person who is careless ahead. In the proposed method, the traffic light is detected employing the images captured by a camera attached to the head of a person. The Histogram of Oriented Gradients(HOG) feature has been adapted to identify and describe a traffic light. The traffic light is recognized by a Random Forest classifier. Then a pattern analysis is performed with the traffic light in the image sequence, and the possibility is judged if one may start to cross the initial part of the pedestrian crossing or one should wait for some time.

The paper also describes a way to find a walkable area on the sidewalk employing the GBS and a way to approach a specified pedestrian crossing employing the GPS signal.

2. Traffic Light Detection

A traffic light is detected by the following procedure; extraction of candidate areas of a target object, feature description of the object using HOG features [7-12], and its detection by a Random Forest classifier designed by learning.

2.1. *Extraction of candidate traffic light areas*

In the proposed method, instead of searching for a traffic light with a whole input image, the search is conducted in the upper half of the image: The reason for this is that a traffic light is normally in the upper half of the images, which a MY VISION system provide. The image is converted to an HSV expression, which is close to human color recognition [12]. An area that might have a traffic light can be searched for by identifying the colors of the light, namely red or green.

Once some candidate traffic light areas are identified, their colored areas receive an expansion processing since they become contracted and are smaller than their original areas. After the expansion, the expanded areas are given labels, and those noisy areas are removed using a certain threshold value.

2.2. *Feature description and recognition*

The HOG feature [7,12] is used in the proposed method to describe the feature of a traffic light, as it is an effective feature for object recognition. The gradient direction is quantized into four bins. This is sufficient since a traffic light mainly has horizontal and vertical line segments and some slant line segments composing a human standing and walking figures. It also contributes to absorbing noisy smaller rotations of MY VISION images caused by a user's natural walk.

For the recognition, on the other hand, Random Forest [13] is employed as a classifier since it can learn and identify multiple target objects. Thus, this classifier will, in the future, contribute to recognizing multiple objects related to road walking, such as pedestrian crossings, traffic lights, various road signs, etc.

3. Judgment on Crossing at a Pedestrian Crossing

The method is explained for analyzing the states of a traffic light and determining the timing of crossing. The analysis is performed based on the hypothesis on pedestrian crossing behavior principle [14].

3.1. *Hypothesis on Pedestrian crossing behavior principle*

The hypothesis on pedestrian crossing behavior principle [14] has been proposed, which gives insight into the relationship between the color of a pedestrian traffic light and his/her behavior.

The following 4 empirical patterns are considered in the hypothesis based on the signal color change of a traffic light.

1. Realizing the change of the light from red to green: The case regarding the duration of the green light is on is observable, and it is considered to be the safest to cross.
2. Realizing the green light display: The case where the length of the green light display cannot be perceived, and it is unknown when it will turn to a flashing signal. The safety or danger to an individual depends on the walking speed of a pedestrian.
3. Realizing the change from green to flashing green: This is the case where the length of the green flashing time is perceptible, which can help a person who is crossing to complete the crossing.

4. Realizing only the flashing green light. The length of the green flashing time cannot be perceived, and it is not recommended to start crossing.

The proposed method identifies the above 4 patterns from the image sequence provided by MY VISION.

3.2. *Discrimination of the state of a traffic light*

This section describes how to discriminate the color and the flashing of a traffic light.

3.2.1. *Color discrimination*

The color of the traffic light is determined based on the HSV model. The color of the signal (red/green) is identified by setting certain thresholds on H, S and V values and counting the number of pixels of each color in the search window.

3.2.2. *Identification of a flashing signal*

The flashing cycle of a traffic light is used to determine whether the signal on a video image frame is flashing or not. The flashing cycle of the traffic light is 0.5 seconds in Japan. If this cycle is satisfied in the frame sequence, it is determined as a flashing signal.

3.3. *Judgment on crossing*

Corresponding to the 4 patterns in the hypothesis on pedestrian crossing behavior principle, there are four signal conditions: (a) red to green, (b) green, (c) green to green flashing, and (d) green flashing. They are identified by examining color and flashing state from the sequence of video frames. The recommendation on the actions are given as follows;

(a) Red to green: One may cross the pedestrian crossing.

(b) Green: One may cross it, if he/she walks normally: Wait till (a) occurs, if he/she cannot walk normally.

(c) Green to green flashing: Wait till condition (a). If one is on the cross, finish crossing.

(d) Green flashing: Wait until condition (a).

It is noted that, in respect to (b), a user's walk speed (normal or slow) is supposed to be taken into account in advance of the method.

The above procedures are illustrated in **Fig.1**.

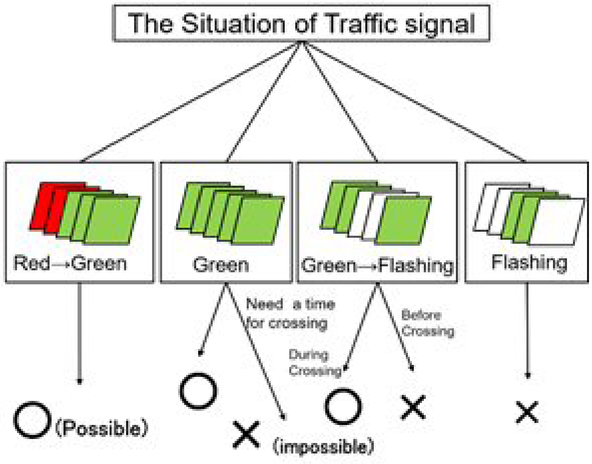

Fig.1. Pattern identification and judgment on crossing: Four patterns are identified from the video and respective actions are recommended.

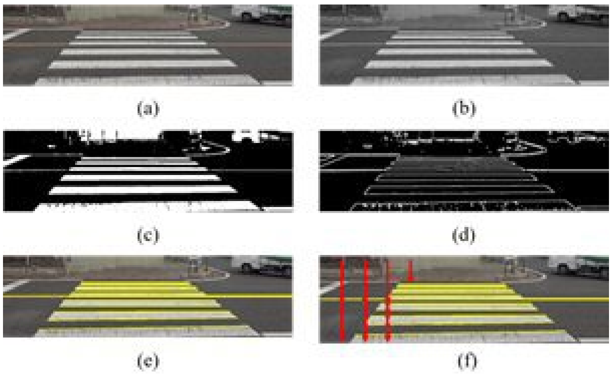

Fig.2. The flow of the pedestrian crossing distance calculation: (a) input image, (b) noise reduction, (c) binarization, (d) edge extraction, (e) line detection and (f) calculation of the distance.

3.4. *Detection of a pedestrian crossing*

When a user is walking over a pedestrian crossing, the proposed method calculates the approximate remaining distance to cross. To do this, the lower half of an input image is processed to extract an image of a pedestrian crossing. The process goes in the following way.

Initially, the image noise is removed by a Gaussian filter (See **Fig. 2(a-b)**). It is then binarized by the discriminant analysis [15] to make it less insensitive to outdoor weather change (Fig. 2(c)). After performing edge extraction with a Sobel filter (Fig. 2(d)), a probabilistic Hough transform [16] (which is faster than normal Hough transform) is applied to the edge image to detect strong horizontal lines composing a pedestrian crossing (Fig. 2(e)). Finally, the number of horizontal lines is counted (Fig.2(f)) by repeating the vertical scan horizontally. Its median value is adopted as the final number of lines on the image. The remaining distance is calculated using the number and considering that the interval between successive lines is about 0.5 meter [17].

The suggestion as to proceed or wait to the user is done by a synthetic voice such as "The traffic signal is red," "The traffic signal will change soon. About 3 meters remaining," etc.

4. Access to a Pedestrian Crossing

The proposed method also guides a user walking on the sidewalk to a specified pedestrian crossing. The strategy is described in the following.

4.1 *Navigation on the sidewalk*

The proposed method finds a walkable area on the sidewalk by avoiding obstacles such as a standing person, a parked bicycle, a puddle, etc. For this purpose, Graph-Based Segmentation (GBS) [18] is applied to an input image, and the extracted area containing a user's foot area is chosen as a candidate walkable area. This procedure is applied to N successive frames, and the area containing the pixels that are the members of all the N candidate walkable areas is finally chosen as a walkable area [19]. The shortest path on it is found employing the A* algorithm [20] and this information is informed to the user.

The above procedure is iterated until a user comes to a specified pedestrian crossing.

4.2 *Access to a specified pedestrian crossing*

A GPS signal is employed for the self-positioning of a user. As the GPS signal value at the specified pedestrian crossing is known in advance, the distance from the user to the pedestrian crossing is evaluated for the user to reach the pedestrian crossing. Since the method introduces numerical evaluation in approaching the destination, it is more exact and stable than the method[19] employing Bags of Features.

4.3 *Facing a pedestrian crossing*

Once a user comes to a pedestrian crossing, the user needs to face it by adjusting his/her body in the right direction before the proposed method starts detecting a traffic light. The method employs Bags of Features (BoFs) [21] to perform this procedure. In this practice, visual words are defined from collected images from various directions at a pedestrian crossing. Then the visual words define BoFs of the images of respective directions. An image fed from an user's camera is examined in regard to its similarity with these BoFs. If the most similar BoFs represent an image facing the pedestrian crossing, the user is standing correctly at the crossing.

5. Experimental Results

An experiment was conducted to examine the performance of the proposed method. The camera is Sony FDR-X3000. It is mounted above the right ear of a subject. Five outdoor videos were taken to evaluate the proposed method. The evaluation was done on (i) traffic light detection, (ii) identification of the state of a traffic light, (iii) judgment on wait/cross action, (iv) calculation of remaining distance, (v) walkable area detection, (vi) self-positioning of a user and (vii) facing a pedestrian crossing.

The evaluation was done employing *recall*, *precision* and *F-value* with respect to (i), (ii), (v) and (vii). With respect to (iii), the evaluation was done by the average accuracy (= the number of true positive frames divided by the number of all frames): With (iv), it was done by the average relative distance errors: With (vi), by the average absolute distance errors.

The results are given in the following. The average F-value of (i) was 0.89. The average F-values of (ii) red, blue, and flashing signals detection were 0.87, 0.89, and 0.94, respectively. The average accuracy for (iii) was 0.91, and the average relative distance errors of (iv) was 25.4%. Furthermore, the average F-value of (v) was 0.86. The average absolute distance errors with (vi) were 2.3m. The average F-value of (vii) was 0.62.

Figure 3 shows some experimental results. The average processing time of the overall process (*navigation on the sidewalk* → *access to a specified pedestrian crossing* → *facing a pedestrian crossing* → *traffic light detection* → *judgement on crossing at a pedestrian crossing*→*detection of a pedestrian crossing*) is approximately 0.99 s/frame by a PC having a i7-4770K, 3.5 GHz CPU.

6. Discussion

A method of navigating a visually impaired person was proposed. By analyzing MY VISION videos, the method navigates a user walking a sidewalk to a specified pedestrian crossing and recommends the user some actions for safe crossing. The recommendation is based on the hypothesis on pedestrian crossing behavior principle [14] and the analysis of the state of traffic light signals.

Expecting results were obtained with (i) traffic light detection using HOG features and Random Forest, (ii) the state (red, green, flash) analysis of the traffic lights, and (iii) judgment on wait/cross actions based on the state analysis. Selecting 4 bins in HOG feature description was effective since it could absorb a varying inclination of the line segments composing a traffic light. On the other hand, (iv) calculation of the remaining distance needs to increase the accuracy. The information on how long one should walk further may be useful in particular to those visually impaired.

Although the evaluation was done to the experimental results on procedures (v), (vi), and (vii), they are based on a rather small number of data. Further experiments are needed to obtain more reliable evaluation results.

The detection of other objects includes road signs for pedestrians and obstacles which get in the way such as pedestrians who are standing or walking towards a user, or a bicycle that has been left behind nearby, etc. Thus, the fact that multiple objects need to be recognized is an issue for further study.

The purpose of the present study is to realize the safety of visually impaired when they walk outdoors. A pedestrian crossing can pose a serious challenge for them. There are three important parameters relating to the safety of crossing a road: the 1st, the length of a pedestrian crossing, the 2nd, the average speed of a user, and the 3rd, the remaining time to cross. The former two can be obtained in advance, but the latter one cannot be acquired through the present traffic light system despite its absolute importance. There are traffic lights which do show how many seconds are left before the signal turns from red to green. The introduction of a reverse traffic light which shows how many seconds are left before the signal turns from green to red will surely contribute to realizing the safety of visually impaired people for crossing the pedestrian crossing.

(a)

(b)

(c)

Fig. 3 Experimental results: (a) Red & wait; (b) green & go; (c) green & go.

7. Conclusion

This paper proposed a method of helping a visually impaired person to navigate outdoors. The method guides a visually impaired person walking on the sidewalk to a specified pedestrian crossing. There it detects a pedestrian traffic light employing MY VISION images and analyzes the states of signals to determine the best action on crossing. The effectiveness of the method was verified by experiments. To realize a practical system for a visually impaired person to walk safely in an outdoor environment, it is necessary to detect and analyze not only traffic lights but also other kinds of objects on a road. These issues are a concern for further study.

References

1. Ministry of Health, Labour and Welfare. "Health care and welfare measures for person with disabilities", Annual Health, Labor and Welfare report 2011-2022, P5, 2011.
2. Bourne, R. R. A., et al. "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis", *Lancet Glob Health*, 2017;5(9): e888–e897.
3. S. Abe, N. Hashimoto. "Survey on walking for the visually impaired", *Bulletin of Hachinohe Institute of Technology*, 24 (2005): 81-92. (In Japanese)
4. Imano, Hosokawa, et al. "Research on autonomous movement of guide robots that guide visually impaired people", *The 3rd Report Image Processing for Crossing Crossing Roads*, 2018 Tohoku Area Young Researcher Research Presentation, 2018. (In Japanese)
5. Baba, Sakai et al. "Real-time recognition of pedestrian signals from moving images", *Information Processing Society of Japan 77th National Convention.* (In Japanese)
6. Fairfield, Nathaniel, and C. Urmson. "Traffic light mapping and detection". *2011 IEEE International Conference on Robotics and Automation (ICRA)*, pp.5421-5426, 2011.
7. N. Dalal, B. Triggs. "Histograms of oriented gradients for human detection", *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Vol.1, pp.886-893, 2005.
8. Q. Zhu, S. Avidari, M-C. Yeh, K-T. Cheng. "Fast human detection using cascade of histogram of oriented gradients", *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, 8pages(IEEE Xplore), pp.1491-1498, 2006.
9. *H-X Jia, Y-J. Zhang. "Fast human detection by boosting histogram of oriented gradients", Proc. of Fourth International Conference on Image and Graphics*, pp.683-688, 2007.
10. Y. Nakashima, J. K. Tan, H. Kim, S. Ishikawa. "A pedestrian detection method using the extension of the hog feature", *Proc. of 2014 Joint 7th International Conference on Soft Computing and Intelligent Systems*, pp.1198-1202, 2014.
11. *T. Lee, M. Hwangbo, T. Alan, O. Tickoo, R. Iyer*. "*Low-complexity hog for efficient video saliency*", *Proc. 2015 IEEE International Conference on Image Processing*, pp.3749-3752, 2015.
12. T. Kumano, J. K. Tan, H. Kim, S. Ishikawa. "Traffic signs and signals detection employing the My Vision system for a visually impaired person", *International Journal of Research and Surveys, ICIC Express Letters, partB: Applications*, Vol.7, No.2, pp.385-391, 2016.
13. L. Breman. "Random forests", *Machine Learning*, Vol.45, No.1, pp.5-32, 2001.
14. K. Hatoyama , S. Sugimori. "Analysis on pedestrian behavior considering spaces before the crosswalk", *Proc. of Infrastructure Planning*, Vol.30, p.171, 2004. (in Japanese)
15. N. Otsu. "A threshold selection method from gray-level histograms", *IEEE Transactions on systems, man, and cybernetics*, Vol.9, No.1, pp.62-66, 1979.
16. N. Kiryati, Y. Eldar, A. M. Bruckstein. "A probabilistic Hough transform". *Pattern Recognition,* Vol.24, No.4, pp.303-316, 1991.
17. Prime Minister's Office, Ministry of Construction: Orders on Road signs, Bound lines and Road markings, No. 3, 1960.
18. P. Felzenszwalb, D. Huttenlocher. "Efficient Graph-based image segmentation" International journal of Computer Vision, Vol.59, No.2, pp.167-181, 2004.
19. J. K. Tan, T. Ishimine, S. Arimasu. "Walk environment analysis using MY VISION: Toward a navigation system providing visual assistance", IJICIC , Vol.15, No.3, 2019.
20. H. Peter, N. Nilsson, B. Raphael. "A formal basis for the heuristic determination of minimum cost paths", IEEE transactions on Systems Science and Cybernetics, Vol.4, No.2, pp.100-107, 1968.
21. G. Csurka, C. Dance, F. Lixin, J. Willamowski, C.Bray:."Visual categorization with bags of keypoints", In Workshop on statistical learning in computer vision, ECCV, Vol.1, No.1- 22, pp. 1-2, 2004.

Authors Introduction

Prof. Dr. Joo Kooi Tan

She is currently with Department of Mechanical and Control Engineering, Kyushu Institute of Technology, as Professor Her current research interests include ego-motion, three-dimensional shape/motion recovery, human detection and its motion analysis from videos. She was awarded SICE Kyushu Branch Young Author's in 1999, the AROB Young Author's Award in 2004, Young Author's Award from IPSJ of Kyushu Branch in 2004 and BMFSA Best Paper Award in 2008, 2010, 2013 and 2015. She is a member of IEEE, The Information Processing Society, The Institute of Electronics, Information and Communication Engineers, and The Biomedical Fuzzy Systems Association of Japan.

Mr. Kohei Kitagawa

He received B.E.and M.E. in Control Engineering from the Graduate School of Engineering, Kyushu Institute of Technology, Japan. His research includes computer vision, pattern recognition, pedestrian crossing detection and analysis.