A Support System for a Visually Impaired Person Finding Bus Route Numbers Employing MY VISION

Daichi Nanaura
Graduate School of Engineering, Kyushu Institute of Technology, 1-1 Sensuicho, Tobata-ku, Kitakyushu, 804-8550, Japan

Seiji Ishikawa[1*], Yui Tanjo[1]
[1]*Faculty of Engineering, Kyushu Institute of Technology, 1-1 Sensuicho, Tobata-ku, Kitakyushu, 804-8550, Japan*
**Emeritus Professor of Kyushu Institute of Technology*
Email: nanaura.daichi608 @mail.kyutech.jp, tanjo@cntl.kyutech.ac.jp

Abstract

A bus is not a very convenient means for a visually impaired person because of the difficulty in identifying its route number, although it is an economical tool for travel. This paper proposes a method of detecting and identifying a bus route number using the MYVISION system which employs an ego camera worn by a visually impaired person. The method finds a frontal area of a bus approaching a bus stop using the video which the camera provides by using optical flow and random forest employing Haar-like features. It then extracts the upper destination panel area followed by the detection of a route number at the right-hand side of the panel. Finally, the detected route number is identified by template matching. In the experiment, various kinds of videos containing the buses of a bus company were captured at different places and different weather conditions, and the effectiveness of the proposed method was shown. The method is now planned to be applied to the busses of various bus companies.

Keywords: MY VISION, visually impaired person, bus route numbers recognition, Haar-like feature, random forest

1. Introduction

According to 2019 WHO report [1], it is estimated that there are at least 2.2 billion people affected by the conditions that cause severe visual impairment worldwide. People with visual impairment can reach daily familiar places without difficulty, but they have difficulty in unfamiliar places. In particular, travelling by bus is reported to be difficult, because identifying the bus route number is not easy for them [2].

Various studies and developments have been conducted worldwide to address this issue. For example, there is a research on having drivers carry special devices to communicate wirelessly with those people who have difficulty in seeing [3], and a research on using GPS and RFID to obtain bus information [4]. However, they require devices to be installed at all buses and bus stops, which is very costly. Another study [5] uses a mobile application and a smartphone camera to identify buses. Although the cost is lower, since the device only needs to be implemented by the user, he/she cannot operate the smartphone unless the user has a certain level of vision

This paper proposes a method that detects a bus from the images of a single camera MY VISION: (a Magic eYe of a Visually Impaired for Safety and Independent actiON) worn by a person with visual difficulties by video image processing. The proposed procedure is as follows: A small camera is attached to the right arm of the person with visual difficulties and he/she stands at the bus stop. When moving objects such as a bus or a car approaches the bus stop, the method judges which vehicle it is. If the method recognizes the object as a bus, the route number is extracted and identified from the bus frontal image and it is informed to the user The proposed system aims to reduce costs and to improve versatility, which was an issue in the related research [4]. The proposed method uses only a single camera and a computer. In addition, the system can be used regardless of the weakness of the eyesight by automatically performing every procedure from start to recognition without requiring any operation by the user, thereby improving convenience in usage.

The proposed method is improved from the related research [6], which identifies whether nor not a moving vehicle is a bus from the images from a head-mounted camera. The present research detects the movement of a bus or a car when it approaches a bus stop and identifies which one it is. The technique proposed in [6] is applied only to the images taken during the daytime by updating Camshift histogram of each frame to track the movement. This makes the system increase the amount of calculation during execution. On the other hand, a bus number recognition technique proposed in [7] conducts to detect the display on the top of the bus and recognize the bus route number by labelling. However, this method can only be employed in the evening when the contrast of the display is large by the LEDs used in the display.

In order to solve the recognition problem of both a bus and its route number, the proposed method combines the problem into a single process and performs the update of the Camshift only once rather than performing it in all frames. Moreover, the proposed method improves the possibility of recognition process even in the daytime when the sunlight is strong using gamma correction to adjust the contrast.

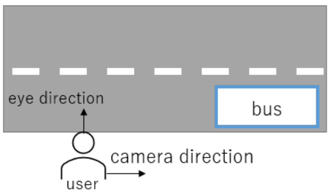
2. Proposed Method

2.1. *Detection of moving objects*

The proposed method replaces the eyes of visually impaired persons with MY VISION. When waiting at a bus stop, a visually impaired person's body faces the road, and the bus approaches from his/her right/left-hand side. Therefore, a MY VISION camera is worn on his/her right upper arm to capture video. A photo of the actual wearing and positioning of a camera is shown in Fig. 1.

The Harris corner detector [8] is applied to the input images to extract feature points. The extracted feature points are tracked using the pyramidal L-K method [9] to acquire optical flows. The acquired optical flows are then clustered using the *k*-means method [10] to extract the candidate regions of moving objects.

(a)



(b)

Fig.1. Photographic environment: (a) A camera mounted on the right arm, (b) location of a user and the bus.

2.2. *Recognition of a bus*

Moving objects in the extracted candidate area described in section 2.1 are then recognized as a bus or other objects. Random Forest [11] is used for the recognition, which is trained using a variety of decision trees. Decision trees with low correlation are created by randomly selecting features at the non-terminal nodes of each decision tree. Learning and testing are carried out using the various decision trees created in this way. Haar-like features [12] are calculated from each training image to make teaching data of each class and a set of feature samples. From this set of samples, a random subset is created. Haar-like feature patterns shown in Fig. 2 are calculated by the following equation;

$$H(A,B) = f(A) - f(B) \qquad (1)$$

where $f(*)$ is the average luminance in each region *.The learning parameters of the Random Forest are shown in Table 1. The image is scanned from the top left to bottom right to obtain these six kind of Haar-like features. Since we need to identify the bus route number at a close distance, unlike the conventional technique [6,7], we do not trace a bus from a distant location. Instead, we use Random Forest to identify the bus at a distance. If an image is identified as a bus for more than five successive frames, it is judged that the bus is in front of or approaching the user, and then the identified area is automatically tracked. We use a Continuously Adaptive Mean Shift (CAMSHIFT) [13] algorithm for tracking, the size of the rectangle area is varied according to the size of the target (a bus) to be tracked. Initially, a Hue histogram within the rectangular area (a bounding box) is calculated, and the probability distribution of a hue is used to search for the similarity of the area in the successive frames, and the search area is shifted to the area where the similarity is larger.

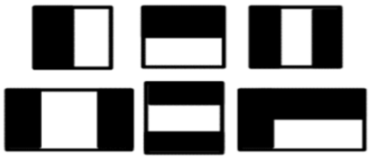


Fig.2. Six Haar-like features used in the method.

2.3. *Application to daytime video*

Table 1 Parameters of Random Forest

data set	Bus ; 50 Background ; 50 Other vehicle ; 50
feature dimensions	5157
Data set size[pixel]	80*48
subset	40
tree	5
Feature selection count	100
Threshold selection count	100

When photographing an LED display of a bus showing the destination and the route number, the camera shutter must be released at a frequency lower than the LED light emission frequency, so that the images do not become dark or do not being reflected. However, if the shutter speed is lowered, the exposure time increases when photographing during daytime, which makes the image overexposed. Overexposure causes the decrease of the recognition accuracy and tracing ability by Camshift. To solve this problem, rather than adjusting the shutter speed, gamma correction is applied to the obtained image. This makes the Haar-like features, *i.e.*, the difference of the brightness values remarkably enhance, and also prevents the color probability distribution from abnormal expansion when applying Camshift to the image for tracing a moving object.

Since the bus is recognized in the distance and moves from the right to the left on the captured video, there is no need to refer to the area on the right of the coordinates that were recognized immediately before.
In addition, since the bus approaches the bus stop at a lower speed and decelerates, the area to be referred can be expanded only on the left-side. The expansion speed of the area to calculate the color probability distribution is determined from the value calculated by assuming the speed range of the bus. This makes the tracing area by Camshift enable to prevent from expanding excessively.

2.4. *Extracting the image of the route number*

In the extraction step of the route number, instead of searching the whole image, we consider and cut out the upper 30% of a rectangular area from the image, and apply HSV transform to the rectangle image. As the route number characteristic is an orange color, we apply the Hue histogram to the rectangle area to obtain a new rectangular area with the orange color distribution. The rectangular area of the bus route number is binarized using the discriminant analysis. Finally, the area of the number is extracted by vertical and horizontal projection of the binary pixels. Template matching is performed to the extracted number to identify it. The process of the route number extraction is illustrated in Fig.3.

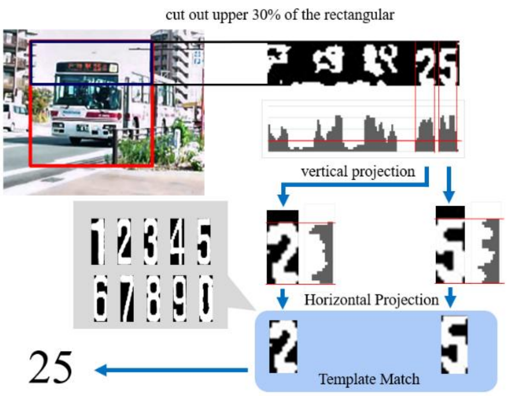


Fig.3. The process of extracting the route number.

3. Experimental Results

3.1. *Bus Recognition*

In the experiment, we used six videos to verify the performance of the proposed method. The evaluation criterion is shown in Fig.4. GT is a ground truth and it stands for the correct area. OA means the detected area. Therefore, the cover value of the GT area with the bus detection area is used for evaluating the bus recognition ability. Here, we set the cover value threshold T to 0.8.

The recognition accuracy (*True*) and false recognition (*False*) are calculated as follows;

$$True = \frac{F_t}{F_{GT}} \times 100\ [\%] \qquad (2)$$

Fig.4. Evaluation criterion of the cover value.

$$False = \frac{F_f}{F_{GT}} \times 100\ [\%] \qquad (3)$$

where, F_{GT} means the total number of Ground Truth frames, F_t means the total number of the frames with cover more than T, and F_f means the total number of the frames with cover less than T.
As a result, the average recognition accuracy (*True*) was 95.5%. The conventional method [6] sets T=0.5, while the proposed method sets T=0.8, which is stricter than the previous method. If an approaching bus cannot be correctly enclosed by a bounding box, the route number extraction becomes difficult.

3.2. *Route numbers detection and recognition*

An experiment was conducted to evaluate the accuracy of the route number recognition with six videos. 'True' was considered correct if the result was correct recognition, while 'undetected' or 'false positive' was judged as False. The evaluation and conditions are shown in Table 2. The average percentage of True responses was 67%. The results of two videos out of six were False. This is because the characters' route number of the video 4 were unclear when binarization was performed to the image. In another video, the route number was not detected. This is due to the proposed route number recognition algorithm is only effective for a two-digit number, and the route number of the experimental video 6 was one-digit number. In this case, we evaluated the result as False. Part of the experimental results are shown in Fig.5.

4. Discussion and Conclusion

A system was proposed to detect a bus and its route number to assist a visually impaired person at a bus stop using MY VISION. Experiments showed the effectiveness of the proposed method, *i.e.*, detection and recognition of busses from moving objects and recognition of the route numbers were successfully done. The accuracy of bus recognition was 95.5% in average, which was improved by about 3% compared to the conventional method [6]. It was also successful in recognizing buses over a wider time range than the conventional method [7]. A more robust algorithm needs to be developed to improve the recognition accuracy of the bus and the detection of route numbers including one-digit numbers.

(a)

(b)

Fig.5. Part of the experimental results: (a) Video 1, (b) video 2.

Table 2 Result of the recognition of route numbers and corresponding time zones.

Videos	Recognition	Time zone
1	True	evening
2	True	daytime
3	True	daytime
4	False	daytime
5	True	daytime
6	False	evening
Accuracy	67[%]	

References

1. WHO: World report on vision, https://www.who.int/publications/i/item/9789241516570, 2019, (Accessed 2024-10-23).
2. Kitakyushu City Disability Welfare Information Center: Hikosen Mirai, https://www.ksjc.jp/index.php?id=2293, (Accessed 2024-8-9)
3. P. Herdiansyah, G. I. Hapsari, G. A. Mutiara, L. Meisaroh: "Smart bus arrival notification for visually impaired", Proceedings of International Conference on Electrical Engineering and Informatics (ICEEI), pp.614-618, 2019.
4. M. Z. H. Noor, I. Ismail and M. F. Saaid: "Bus detection device for the blind using RFID application", International Colloquium on Signal Processing & Its Applications (CSPA), pp. 247-249, 2009.
5. P. Wongta, T. H. Chalidabhongse: "Vision-based bus route number reader for visually impaired travelers", International Conference on Imaging, Signal Processing and Communication (ICISPC), pp. 64-69, 2018.
6. K. Ishitobi, J. K. Tan, H. Kim, S. Ishikawa: "Detection of a specific moving object from head-mounted camera images", 2017 IEEE/SICE International Symposium on System Integration (SII), pp. 817-822, 2017.
7. J. K. Tan, Y. Hamasaki, Y. Zhou, K. Ishitobi: "A method of identifying a public bus route number employing MY VISION", Journal of Robotics, Networking and Artificial Life, Volume 8, Issue 3, pp. 224 -228, 2021.
8. C. Harris, M. Stephens: "A combined corner and edge detector", Proc. of the 4th Alvey Vision Conference, pp.147-151, 1988.
9. J. Y. Bouguet: "Pyramidal implementation of the Lucas Kanade feature tracker description of the algorithm", Intel Corporation, Microprocessor Research Labs, OpenCV Documents, 9 pages, 2000.
10. J. MacQueen: "Some methods for classification and analysis of multivariate observation", Proc. of the 5th Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, pp.281-297, 1967.
11. F. Breiman: "Random forests", Machine Learning, Vol. 45, No. 1, pp.5-32, 2001.
12. P. Viola, M. Jones: "Rapid object detection using a boosted cascade of simple features", Proc. of the Conference on Computer Vision and Pattern Recognition, Vol. 1, pp.511- 518, 2001.
13. G.R. Bradski: "Computer vision face tracking for use in a perceptual user interface," Intel Technology Journal Q' 98, 15pages, 1998.

Authors Introduction

Mr. Daichi Nanaura

He received his Bachelor's degree in Engineering in 2023 from the Faculty of Engineering, Kyushu Institute of technology in Japan. He is currently a master student in Kyushu Institute of Technology, Japan

E. Prof. Dr. Seiji Ishikawa

Seiji Ishikawa graduated from Tokyo University and was awarded BE, ME and PhD there. He is now Emeritus Professor of Kyushu Institute of Technology. He was Visiting Researcher of the University of Sheffield, UK, and Visiting Professor of Utrecht University, NL. His research interests include visual sensing & 3-D shape/motion recovery. He was awarded The Best Paper Awards in 2008, 2010, 2013 and 2015 from BMFSA.

Prof. Dr. Yui Tanjo

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