

## Research Article

# A Method of Identifying a Public Bus Route Number Employing MY VISION

Joo Kooi Tan\*, Yosiki Hamasaki, Ye Zhou, Ishitobi Kazuma

*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 20 August 2021

*Keywords*Bus route number  
moving area detection  
Haar-like filters  
random forest  
the visually impaired  
MY VISION**ABSTRACT**

In response to requests from visually impaired people for better assistance tools in their daily lives and due to the difficulty of using public transportation, it is proposed a bus boarding support system using MY VISION. For the support, a method of identifying a bus route number is proposed in this paper. A bus approaching a bus stop is detected from MY VISION images that utilizes the Lucas–Kanade tracker. A bus route number area is then acquired with the help of the Haar-like filters and the random forest. Finally, the number from the area is extracted and identified through pattern matching. The effectiveness of the proposed method was shown by experiments. The results indicate that it is possible to realize a practical boarding support system for the visually impaired.

© 2021 The Authors. Published by Atlantis Press International B.V.

This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).**1. INTRODUCTION**

According to the statistics survey from the Ministry of Labor in the past 23 years, there are about 310,000 people with visual impairment in Japan [1]. Meanwhile, about 70% of these people are over 65 years old. It can be seen that, as social aging goes, the number of people with visual impairment increases further. For the visually impaired, considering practicality and economy, they usually choose public transportation. However, public transportation, a bus, for example, needs some kinds of visual information, such as the route number and entrance location other than knowing its approach and stop at a bus stop where a user is. Acquisition of the above information is vital for the visually impaired. Therefore, it is necessary to develop a support system that enables visually impaired people to easily use public transport.

In this paper, we propose a bus boarding support system for the visually impaired that detects a bus in service and identifies the bus route number employing MY VISION. MY VISION is an ego-camera system which provides the images from the (virtual) view point of the visually impaired. The proposed system assumes that a user is equipped with an ego-camera on his/her head, and it processes the images provided from the camera to detect a bus and to identify the bus route number.

In the following, related work is described in Section 2, procedures of the proposed system are overviewed in Section 3, a method of detecting an approaching bus from images is addressed in Section 4, a method of finding a bus route number area is explained in Section 5, and an identification strategy of a bus route number is stated in Section 6. Experimental results

are shown in Section 7 and Section 8 provides a discussion and conclusion.

**2. RELATED WORK**

From a review of the literature, it was found that there are many automatic bus number recognition systems. However, most of them use active sensors, such as Global Positioning System (GPS) tracking systems, Radio Frequency Identifier (RFID), etc.

For vision-based methods, Guida et al. [2] and Wongta et al. [3] proposed the detection and identification of bus route numbers. Their method uses Adaboost-based cascaded classifiers to locate the main geometric entities of the bus facade, and the matching is improved by robust geometric matching. Then the bus facade image is converted to a Hue Saturation Value (HSV) color space and threshold operations are applied to segment each number.

Pan et al. [4] proposed another vision-based method that detects information from the front of the bus. Their system can be divided into two subsystems: bus detection and bus route number detection. In the bus detection, they use Histograms of Oriented Gradients (HOG) and Support Vector Machine (SVM) to detect the bus position. In the bus route number detection, they use adjacent character grouping and sophisticated edge detection to find candidate regions. The Haar-like feature is extracted from edge distribution. The acquired features are input into Adaboost to classify each component. Finally, they combine an optical character recognition software with a text-to-speech synthesizer to produce audio.

However, the aforementioned methods are insufficient for practical use in that their computational complexity is high, and the recognition accuracy and the effect of tracking a bus still need to be improved to gain better performance.

\*Corresponding author. Email: [etheltan@cntl.kyutech.ac.jp](mailto:etheltan@cntl.kyutech.ac.jp)

### 3. OVERVIEW OF THE PROPOSED METHOD

The proposed method is separated into three parts; (i) bus area detection, (ii) bus route number area detection, and (iii) tracking of the bus route number area and bus route number identification.

In (i), a moving bus area is detected on a fed image by the method [5] using optical flow analysis. Then, in (ii), a bus route number area is extracted by the use of Haar-like features [6] and random forest [7]. Finally, the bus route number area is tracked on the video by CamShift [8] and, at the same time, the bus route number is identified by pattern matching.

### 4. DETECTION OF A BUS AREA

Firstly, the area containing a bus in service is searched in an input image. The bus detection unit uses the Harris corner detector [9] to extract feature points from the input image. The extracted feature points are tracked by the pyramidal Lucas–Kanade tracker [10]. The optical flows obtained by the above method include those optical flows caused by a moving object (referred to as internal optical flows) and those caused by the motion of the camera (external optical flows). Therefore, the obtained optical flows need to be processed to exclude those caused by camera movement. For this purpose, the projective transformation given by

$$\begin{cases} x_2 = \frac{h_0x_1 + h_1y_1 + h_2}{h_6x_1 + h_7y_1 + 1} \\ y_2 = \frac{h_3x_1 + h_4y_1 + h_5}{h_6x_1 + h_7y_1 + 1} \end{cases} \quad (1)$$

and Random Sample Consensus (RANSAC) [11] are used. Assuming that two successive images are mutually related by the projective transformation, Equation (1) expresses the relation between the point  $(x_1, y_1)$  in the first image and a corresponding point  $(x_2, y_2)$  in the second image. RANSAC is used to determine the parameters  $h_i$  ( $i = 0, 1, \dots, 7$ ) of Equation (1). Those point pairs between the two images which do not satisfy the relation given by Equation (1) are regarded as giving the internal optical flows.

After obtaining the internal optical flows, the points which compose the internal optical flows are clustered by  $k$ -means clustering algorithm [12], and a bus area is finally determined.

### 5. DETECTION OF A BUS ROUTE NUMBER AREA

Once a bus area is obtained in an input image, a bus route number area is searched for from a frontal bus image by a bus route number area detection unit. The unit describes a candidate area employing a Haar-like feature vector, and a random forest classifier judges if the vector represents a bus route number area. Figure 1 shows a bus route number area.

#### 5.1. Haar-like Feature

The feature of the bus route number area is described by the Haar-like feature in the proposed method. In previous methods, the most

commonly used feature for vehicle description is the HOG feature. However, when the shape of two regions is similar, the HOG feature is prone to generate similar histograms, which will affect the subsequent recognition tasks negatively. Therefore, in the proposed method, the Haar-like feature is used for regional characterization. Its main advantage is that it characterizes not only the shape of the object, but also the brightness distribution of the image area effectively. Figure 2 shows six Haar-like feature filters employed in the proposed method. Calculation with the filter is done by the following formula;

$$H(A, B) = F(A) - F(B) \quad (2)$$

Here,  $F(A)$  and  $F(B)$  are the average brightness of a white area  $A$  and a black area  $B$ , respectively.

As shown in Figure 2, the six Haar-like features include two edge features (upper left and middle), three linear features (upper right, and lower left and middle) and the center-around feature (lower right).

#### 5.2. Random Forest

After the Haar-like features of the bus route number area have been obtained, the random forest [7] is used as a classifier for finding the area from an input image. The random forest classifier is chosen, since the proposed method will be expanded in the future to recognize various objects related to the walking outdoors and the random forest is a multi-class classifier. For training the classifier, Haar-like feature vectors are calculated using a number of images of a bus route number area as positive data and a number of arbitrary images as negative data. The image size of a bus route number area is standardized to  $100 \times 30$  pixels taking the shape of the area into account.

The random forest classifier determines a class by feeding the feature vector of an unknown  $100 \times 30$  pixels image into the learned decision trees. The class distribution  $P_t(c/S)$  at tree  $t$  is calculated

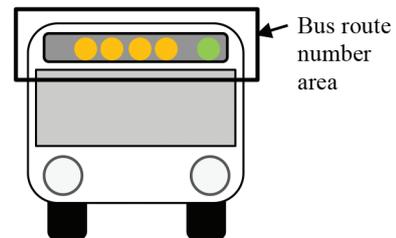


Figure 1 | Bus route number area.

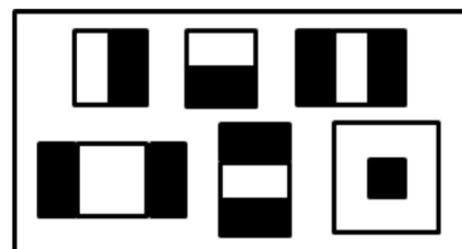


Figure 2 | Six types of Haar-like feature filters.

when the unknown data finally reaches the terminal node of tree  $t$ , and it is averaged among the decision trees by the following equation;

$$P\left(\frac{c}{S}\right) = \frac{1}{T} \sum_{t=1}^T P_t\left(\frac{c}{S}\right) \quad (3)$$

Here,  $T$  is the number of decision trees in the random forest, and  $P(c/S)$  is the average class distribution.

## 6. TRACKING AND IDENTIFICATION

Finally, we describe a tracking and identification unit. Once a bus area is detected by the bus detection unit, the next unit is applied to the bus area to extract a bus route number area. Once this area is extracted, text regions are searched by analyzing the area, and they are processed to identify the route numbers. The bus route number area is then tracked to find itself in the next frame. This process is repeated so that the bus route number is identified all the time.

### 6.1. Tracking of a Bus Route Number Area

The tracking unit uses the CamShift [8] to perform tracking with the color histogram of a bus route number area in a detection window as the initial value. However, due to the fact that the detected target object, i.e. a bus route number area, is gradually approaching an observing camera, the appearance of the object changes gradually. Then the color histogram obtained from the first detection of the target object needs to be updated. In the proposed method, a window slightly larger than the initial tracking window created by the CamShift is set during the tracking. If the target object is detected again in the window, the color histogram for the tracking is updated. Therefore, the histogram of the CamShift used in the method changes with the appearance change of the target object.

### 6.2. Identifying a Bus Route Number

The obtained bus route number area needs further processing to get the text area. According to the fact that the LED display area of a Japanese bus has a fixed position distribution in the bus route number area, a relatively accurate route number position can be obtained. HSV color image processing is employed to get the binary image of the text area, and labeling is done to segment the text area as shown in Figure 3a. The labeled segments are surrounded by a rectangle as shown in Figure 3b, and those rectangles which have the height–width ratio  $r$  ( $0.45 < r < 1.10$  in the experiment) and the size above a certain threshold are chosen. Template matching is then performed on the chosen rectangles to identify the bus route number. Note that the templates for the matching are 10 digits.

## 7. EXPERIMENTAL RESULTS

### 7.1. Set Up

The proposed method was examined in regard to its performance experimentally.

The method analyzes a video provided from a camera worn on the head of a user and taking images of an approaching bus from the sidewalk. Four videos (scenes) were taken and used for the experiment; Scene 1 (67 frames), Scene 2 (103 frames), Scene 3 (86 frames) and Scene 4 (59 frames). The former two scenes were taken on a sunny day, whereas the latter two on a cloudy day.

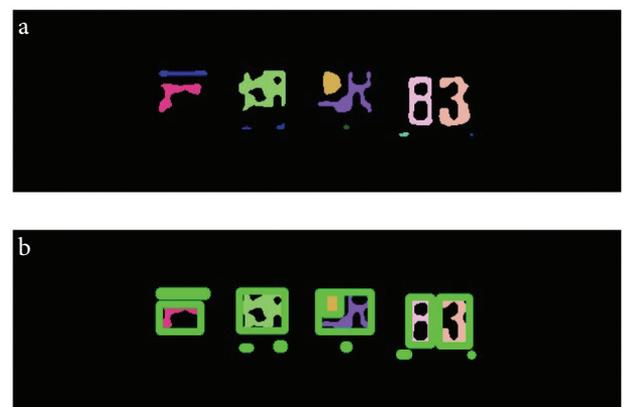
Although the first step of the method is to detect a bus area from a video, this paper concentrates on the detection of a bus route number area and the identification of a bus route number. The results on them are shown in the following tables.

### 7.2. Detection of a Bus Route Number Area

A random forest classifier was trained to detect a bus route number area. Data on the number of used positive/negative images and other parameters are given in Table 1. The performance of the classifier based on the test data is shown in Table 2. Table 3 shows the detection rate when the classifier was applied to the four videos. The dimension of the Haar-like feature vector is 5102 in the experiment.

### 7.3. Identification of a Bus Route Number

Finally, the result on a bus route number identification with respect to the four videos is given in Table 4. Examples of a bus detection, a bus route number area detection and a bus route number identification are depicted in Figure 4, where a green box and a red



**Figure 3** | Analysis of a bus route number area. (a) Binarization and labeling of a text area, (b) rectangle representation of the labeled segments.

**Table 1** | Numerical data on training the random forest classifier

Training images	Positive image: 1400 Negative image: 950
Test images	Positive image: 530 Negative image: 501
Image size	100 × 30 pixels
Number of subsets	30
Tree depth	5

**Table 2** | Performance on the bus route number area detection (%)

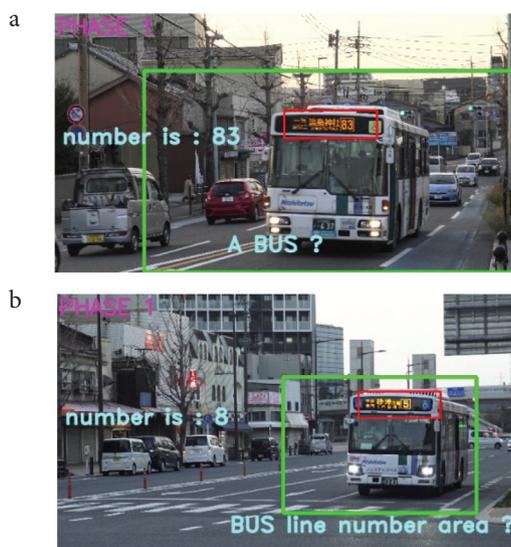
	Bus route no. area	Background
Bus route no. area	93.2	6.8
Background	9.0	91.0

**Table 3** | Detection rate with the four videos (%)

Video no.	Detection rate	Undetected rate	False detection rate
Scene 1	77.6	1.5	20.9
Scene 2	88.4	1.9	9.7
Scene 3	74.4	12.8	12.8
Scene 4	81.4	18.6	0.0
<b>Average</b>	<b>80.5</b>	<b>8.7</b>	<b>10.8</b>

**Table 4** | Result on a bus route number identification (%)

	Identification rate
Scene 1	36.5
Scene 2	69.2
Scene 3	50.0
Scene 4	68.8
<b>Average</b>	<b>56.1</b>

**Figure 4** | Results on bus route number detection. A green box and a red box show a detected bus area and a bus route number area, respectively. The identified bus route numbers, (a) 83 and (b) 8, are given in the middle left of the images.

box indicate a moving object area and a bus route number area, respectively. The bus route number is shown in the middle left of the image. The average processing time is 0.44 s/frame by a PC having a 3.2 GHz Central Processing Unit (CPU).

## 8. DISCUSSION

This paper proposed a method of assisting boarding a public bus for a visually impaired person. It finds a bus approaching a visually

impaired user from a video provided by MY VISION and detects a bus route number area on the front of the bus by a random forest classifier based on Haar-like feature vectors. A bus route number image is then extracted from the area and identified by pattern matching.

Experimental results on the bus detection and on the bus route number area extraction proved the effectiveness of the proposed method. To increase the accuracy of the bus detection further, interference from surrounding moving cars should be excluded, since the cars of the same lane provide the optical flows similar to that of a bus. This causes negative influence to the bus area detection. As for the bus route number area extraction, a larger number of data for training the random forest classifier should be prepared for, so that various lighting conditions may be considered.

On the other hand, results on the identification of a bus route number still needs further improvement. In particular, more exact segmentation of the digits in the bus route number area should be considered under various conditions on illumination and light reflection.

Improving methods of the segmentation and the digits identification remain for further study.

A method [5] of detecting a public bus approaching a bus stop and finding an entrance of the bus stopped in front of a visually impaired user has been proposed. The method proposed in this paper enhances the method [5] by adding the function of a bus route number identification to it. The overall system will lead to the realization of a useful bus boarding assistance system for the visually impaired.

## 9. CONCLUSION

This paper proposed a method of identifying a public bus route number using MY VISION. The method includes three main stages; detection of a bus, detection of a bus route number area, and identification of a bus route number. Experimental results show satisfactory performance with the first two stages, whereas the last stage needs improvement in text segmentation and digits identification. This remains for further study.

The method is expected to offer one of the useful bus boarding support tools for the visually impaired people in the future.

## CONFLICTS OF INTEREST

The authors declare they have no conflicts of interest.

## REFERENCES

- [1] Ministry of Health, Labor and Welfare, Social Assistance Bureau, Disability Health and Welfare Department, 2011 Survey on Difficulty in Living (Fact-finding Survey on Children at Home, etc.) Results, Ministry of Health, Labor and Welfare, Social Assistance Bureau, Disability Health and Welfare Department, 2013, p. 5.
- [2] C. Guida, D. Comanducci, C. Colombo, Automatic Bus Line Number Localization and Recognition on Mobile Phones—A

- Computer Vision Aid for the Visually Impaired, in: G. Maino, G.L. Foresti (Eds.), International Conference on Image Analysis and Processing (ICIAP), Springer, Berlin, Heidelberg, 2011, pp. 323–332.
- [3] P. Wongta, T. Kobchaisawat, T.H. Chalidabhongse, An automatic bus route number recognition, 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), IEEE, Khon Kaen, Thailand, 2016, pp. 1–6.
- [4] H. Pan, C. Yi, Y. Tian, A primary travelling assistant system of bus detection and recognition for visually impaired people, 2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), IEEE, San Jose, CA, USA, 2013, pp. 1–6.
- [5] 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), IEEE, Taipei, 2017, pp. 817–822.
- [6] H. Gong, L. Chen, C. Li, J. Zeng, X. Tao, Y. Wang, Online tracking and relocation based on a new rotation-invariant Haar-like statistical descriptor in endoscopic examination, IEEE Access 8 (2020), 101867–101883.
- [7] L. Breiman, Random forests, Mach. Learn. 45 (2001), 5–32.
- [8] O.D. Nouar, G. Ali, C. Raphael, Improved Object Tracking with Camshift Algorithm, Proceedings of the 2006 IEEE International Conference on Acoustics Speech and Signal Processing, IEEE, Toulouse, France, 2006, pp. II-657–II-660.
- [9] C. Harris, M. Stephens, A Combined Corner and Edge Detector, in: C.J. Taylor (Ed.), Proceedings of the 4th Alvey Vision Conference, The Plessey Company plc, University of Manchester, England, 1988, pp. 147–151.
- [10] J.Y. Bouguet, Pyramidal implementation of the Lucas Kanade feature tracker description of the algorithm, Intel Corporation, Microprocessor Research Labs, OpenCV Documents, 2000, p. 9.
- [11] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (1981), 381–395.
- [12] J.A. Hartigan, M.A. Wong. Algorithm AS 136: a K-means clustering algorithm, J. R. Stat. Soc. 28 (1979), 100–108.

## 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.

### Ms. Ye Zhou



She received B.E. in mechanical design and manufacturing and automation from Yangzhou University, China and M.E. in Control Engineering from the Graduate School of Engineering, Kyushu Institute of Technology, Japan. Her research includes computer vision, bus line number detection and recognition.

### Mr. Yosiki Hamasaki



He received B.E. in Control Engineering from the Graduate School of Engineering, Kyushu Institute of Technology, Japan. His research includes image processing, and pattern recognition.

### Mr. Ishitobi Kazuma



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, and object tracking.