International Journal of New Technology and Research (IJNTR)
ISSN:2454-4116, Volume-4, Issue-9, September 2018 Pages 43-52

Classification of Viewpoints Related to Bus-Waiting for the Assistance of Blind People

Watcharin Tangsuksant, Chikamune Wada

***Abstract*—Considering the current smartphone technology, blind individuals cannot detect oncoming buses from their smartphones. Consequently, this paper proposes the viewpoint classification for blind individuals during situations of non-congested traffic. The definition of ideal viewpoints for this research depends on two main factors, which are the suitable tilt and suitable panning of a smartphone by which it is possible to detect and recognize the bus number. This research aims to find the suitable panning of a smartphone, which related to suitable viewpoint, using the combined simplicity of image processing. The proposed concept for interpreting viewpoints is to extract essential features from the road area via images. These features comprise road length, vanishing point, and the percentage of road area. For this research, various roadside scenarios were tested under both daytime and nighttime illumination. The experimental outcome shows the successful classification of viewpoints with regard to the requirement of blind persons waiting for a bus on the roadside. These results were found to be robust and feasible for roadside scenes under various illumination conditions.**

***Index Terms*—blind persons, road-feature extraction, road segmentation, viewpoint classification.**

I. Introduction

According to the Ministry of Social Development and Human Security of Thailand, 2017, the official number of visually impaired people was 186,639 [1]. However, many visually impaired people yet remain undocumented. An estimated 680,000 unofficial visually impaired people—including those who are completely blind—live in Thailand, according to the Thailand Association of Blind website [2]. Unequivocally, blind people have to face many problems in their daily life, such as reading traffic signs and warning labels in public places. Moreover, travel is also a major concern because they cannot sense the obstacles in front of them. However, these problems might be overcome in developed countries where an assistive system for visually impaired people is available. For example, braille is imprinted on locations such as stairway railings indicating the way toward the exits and entrances as well as toward individual platforms in train stations [3]. Furthermore, automatic voice warning is provided in some dangerous places: almost intersections in Japan, for instance, usually incorporates traffic-light sounds to convey the status of traffic lights to blind individuals [4]. In contrast, no well-designed systems exist in many developing and underdeveloped countries including Thailand.

Herein, we would like to focus on a bus service for blind people who travel independently. In Thailand, because of poor public transportation systems, there are no systems to support disabled people waiting for buses, and the estimation of bus arrival times is impossible for everyone. There are no announcements regarding the details of bus routes and upcoming buses either; therefore, users have to observe the bus number by themselves at the bus stop, and then wave their hand to signal the bus driver to pick them up. Sometimes, blind individuals who travel independently can obtain help from neighboring people at the bus stop. Unfortunately, blind individuals often face extreme difficulty when they stay alone or have no one to assist them at the bus stop.

Some previous research has focused on bus waiting for blind individuals. Existing bus identification methods for blind individuals can be categorized into two main systems. The first system uses transceiver communication [5-11] such as RFID, Bluetooth, WiFi or Global Positioning System (GPS). The general concept of this technique is the installation of a receiver-transmitter with the user's device, such as Personal Digital Assistant (PDA) [6-7] or smartphone [5, 10-11]. For example, E. A. B. Santos proposed [5] the wireless interactive system in which the frequency of 2.4 GHz. was used for communication via each installed module between the user, the bus station, and the bus. Users can select the bus line on the smartphone and then the bus station module will receive the requirement from a user module when checking if the bus line is correct. After, the module at the bus station will send the data to the bus module to notify the bus drivers on the bus panel. A similar bus identification system was presented using RFID module [6, 8-9]. Moreover, the VIABUS [11] is a smartphone application in Thailand that applies the GPS module to search the arrival bus. However, it can only search the buses that had the GPS module installed. Although a transceiver communication system is a good idea to identify the bus for visually impaired people, this communication system is quite large for practical use because transceiver modules have to be installed for every bus station, buses, and user. Further, system maintenance is difficult when some part of the module is not working.

On the other hand, the second system proposed is based on an image processing technique [12-16] to capture the bus numbers of the oncoming buses. Image processing is a fascinating technique that can be widely used for several applications, including bus identification for visually impaired people and blind individuals. For example, Pitchakorn Wongta *et al.* [12], in 2016, proposed an automatic bus route number recognition system in Thailand.

Watcharin Tangsuksant, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Hibikino 2-4 Wakamatsu-ku Kitakyushu Fukuoka 808-0196, Japan
Chikamune Wada, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Hibikino 2-4 Wakamatsu-ku Kitakyushu Fukuoka 808-0196, Japan

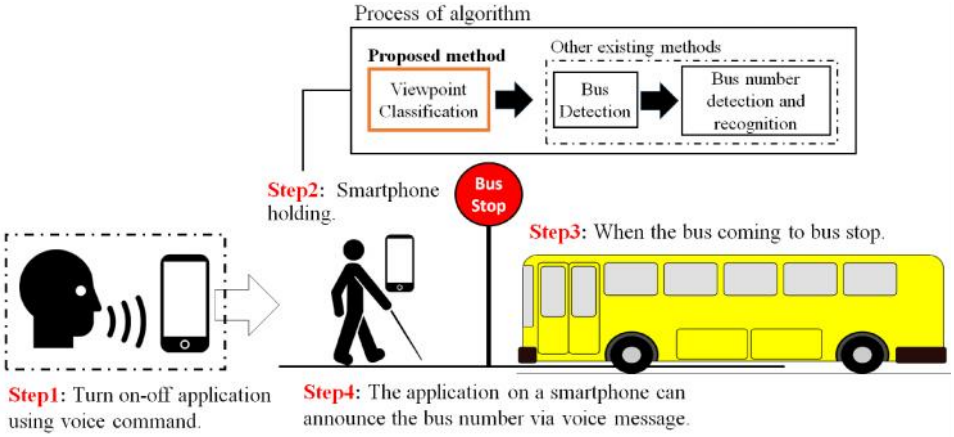

Fig.1 Overview of system design, other existing methods, and proposed method.

Their experimental results showed an accuracy of 73.47% for digit segmentation that the researchers themselves collected, that is, the Thai bus dataset. Furthermore, they interviewed some Thai individuals with poor vision concerning the difficulty they faced when traveling independently using public transportation. In the interview, they mentioned reading the bus number is still challenging for them, although they can see a blurry oncoming bus. The problem of bus number recognition for blind individuals has not only drawn the attention of Thai researchers, but also has also piqued the interest of several researchers from other countries such as Taiwan, Korea, Italy, and the U.S. [13–16].

Based on the above-mentioned systems, from the viewpoint of ease of use for blind people, we proposed our own system shown in Figure 1. Firstly, a voice command is used for turning the application on or off (Step 1). Subsequently, users use their smartphone to take a video when waiting for an oncoming bus at the bus stop (Step 2). When a bus comes to the bus stop (Step 3), the application will convey the bus number via a voice announcement to the user (Step 4). This way, users can better board the correct bus independently [12–16]. However, the process of algorithms in existing research solely focused on two main steps, which were bus detection and bus number recognition as shown in step 2 of Figure 1. In fact, when the blind individuals freely held the smartphone to capture a video, they could not know the viewpoints they took. Further, various viewpoints can appear on a smartphone screen, and some viewpoints might be unsuitable to capture the oncoming bus. For example, if they are standing behind an obstacle, such as an electricity pole or other big obstacle on the roadside. In this case, they may obtain an image of the whole or part of the obstacle instead of the oncoming bus. Even in ideal situations without obstacles, it is still difficult to inform users how suitable or unsuitable their viewpoints are since it can be impossible to capture the bus on every frame of their video. For these reasons, other existing methods could not perform the system completely.

There is no previous research involving the viewpoint classification for aiding blind individuals. The final goal of our application is to let the blind know how to adjust the camera in order to obtain a suitable viewpoint for bus number detection. Figure 2 shows the main step of a whole viewpoint classified algorithm that consists of four important processes. Firstly, the system acquires video. Second, the process of

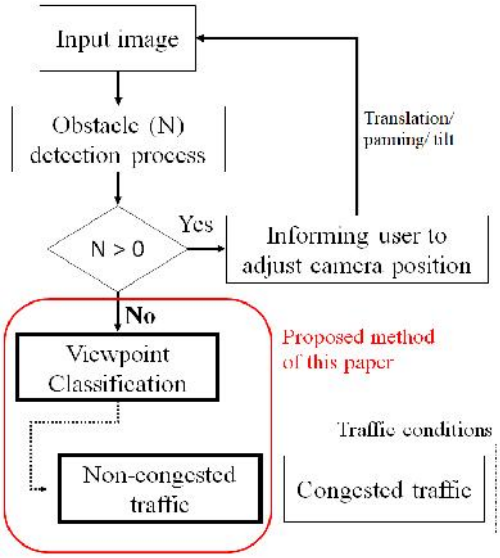

Fig. 2 Main step of viewpoint classified algorithm and proposed method of this paper.

obstacle detection is performed. In case of obstacles appearing, the system will inform the user to adjust the camera position, such as translation, panning the camera leftward/rightward, and tilting the camera upward/downward. The viewpoint classification will be started when the obstacles do not appear in the image. However, the situation on the road can distinguish between two main conditions that consist of congested and non-congested traffic. In order to achieve the final goal of our application, this paper proposes a viewpoint classification for non-congested traffic, as shown in the red block of Figure 2. In addition, the combined simplicity of image processing and the relatively light load calculation techniques are used in conjunction with daily-use systems, such as smartphones, without adding any special equipment.

II. Suitable Viewpoints Definition for Bus Waiting

There is no previous research that proposes the viewpoint classification for blind individuals waiting at a roadside bus stop. Suitable viewpoints should be defined before explaining other sections. When taking the photo, various viewpoints that depend on the two main factors of panning and tilt of camera appear. In Figure 3, an example of the relation between the camera's position (tilt and panning) and image viewpoints of the roadside at the bus stop is shown. For instance, in Image view 1, Image view 2, and Image view 3,

the unsuitable viewpoints are shown. Image view 1 was taken by a very low-angle shot and with unsuitable panning because it is impossible to detect the bus number with this viewpoint. Further, Image view 2 was taken with a suitable tilt, but unsuitable panning because the route bus number was too close to the user. Image view 3 shows the suitable panning of camera, but it is taken by a very high-angle shot that resulted in a large portion of sky appearing in the image. In the suitable instance of a viewpoint for bus waiting on the roadside, as shown in image view 4, the image was taken using suitable tilt and panning. Although humans can visualize and understand those viewpoints in Figure 3 automatically, it is difficult for a machine or computer to interpret the meaning of suitable camera tilt and panning.

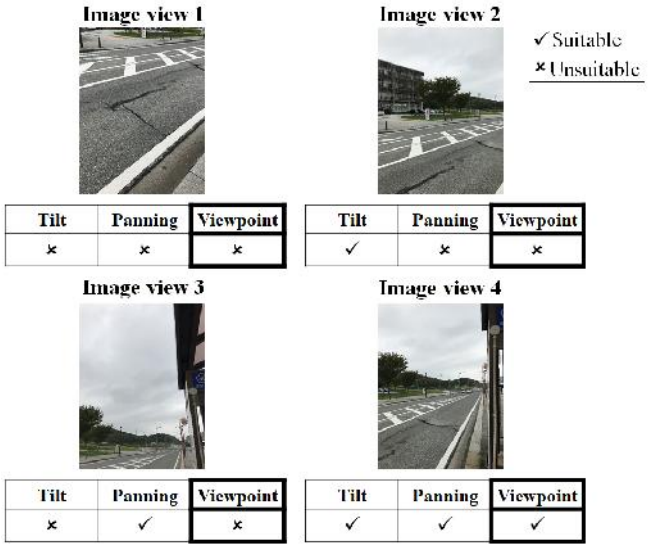

Fig. 3 Example of the camera's position and their image views.

The suitable camera tilt is achieved when the smartphone camera is held vertically. However, the suitable camera panning definition is more complicated to interpret because we have to consider the actual image view and whether it is feasible and suitable for route bus number recognition. For a suitable direction of the image to be used to recognize the correct bus, the route bus number detection and recognition are important for effective directional investigation. In order to estimate the image without the oncoming bus, this paper presents a new method for bus position estimation and suitable camera panning.

First, the image viewpoint that shows the oncoming bus will be considered as shown in Figure 4a in order to calculate the horizontal line (L) between the vanishing point and the bus position. The first step for defining suitable camera panning is to estimate the possible longest distance for bus number recognition. The route bus number (0–9), some English alphabets (A and B), and some Thai alphabets (ก ต ป อ), which include the necessary characters of route bus number of Thailand, are simulated by a 15 cm × 15 cm board that is the same as the standard size of the route bus number sign in Thailand; the image size used is 800 px × 600 px. For this testing, those simulated numbers and alphabets are tested in the ideal conditions of daylight (210.00–230.00 lux) and nightlight (0.80–1.00 lux). Then, different distances between the camera and those characters are tested for recognition. According to the experimental conditions, a distance of 15 meters is selected from the starting distance, within which the system can recognize the route number of the bus, as the reference distance. After the longest distance for the route bus number recognition has been estimated, the next step is capture a photo of the oncoming bus at the bus stop within the reference distance, as shown in Fig. 4a. The distance is measured from the standing point of the user to the façade of the oncoming bus. In order to estimate the image viewpoint without the oncoming bus, the vanishing point (Vp) of the perspective image is defined by two convergent lines of the white line marker on the edge of the road. After, an L horizontal line is drawn from Vp to the square boundary of the detected bus, as shown in Figure 4a.

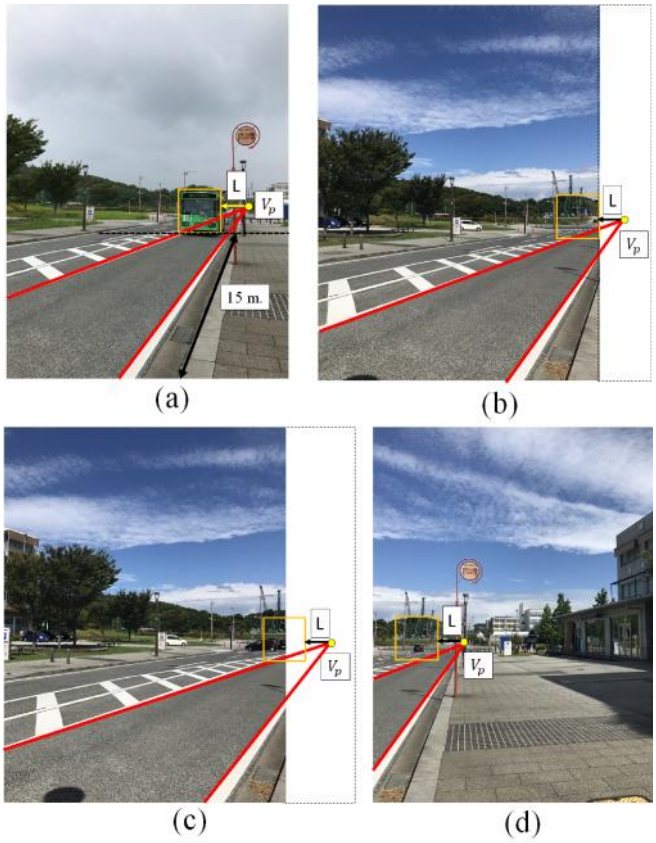

Fig. 4 (a) Oncoming bus at a distance of 15 meters from the camera, (b), (c), and (d) Viewpoint estimation without oncoming bus using vanishing point and L line.

L line is measured as 8.70% (0.49 in) compared with the width of the original size of the image (5.63 in), as shown in Figure 4a. In case there is no bus on the road, the L line can be used to estimate the bus's position at the longest distance for route bus number recognition. For example, Figure 4b shows the suitable camera panning. By marking the vanishing point and drawing the L line, we can estimate the position of the bus in the image where the route bus number will appear when the bus arrives at that position. On the other hand, when the camera pans lightly, as shown in Figure 4c, the bus position can be estimated, but appears with just half of its façade. Thus, it can be claimed that the image viewpoint is unsuitable for the bus number recognition because the route bus number cannot be recognized at the reference distance. Although the bus number can be recognized in Figure 4d at the reference distance, the moving bus will be seen in the image frame only

for a short time. Therefore, it would be better if the camera were panned to the left side in order to see the bus for a longer period of time. Actually, the image viewpoint in Figure 4b is the perfect camera panning to recognize the route number of the bus because the bus number can be recognized at the farthest right side of image; this is the first position in the image viewpoint to see the bus. However, it will be extremely impractical to define only a single suitable viewpoint for the users to hold their smartphone. Consequently, this paper defines the suitable viewpoint within 25% of the range between V_p and the farthest left side in the perfect camera direction that the vanishing point of perspective images have to fall in, as shown in Figure 5a. However, setting a value of 25% is an assumption of the proposed application. Furthermore, there are differences of V line, H line length, and road area between the suitable and unsuitable camera panning as shown in Figure 5a and Figure 5b. For example, the length of V line in Figure 5b is longer than that in Figure 5a, and the length of the H line does not show in Figure 5b, but it shows the long length in Figure 5a. Therefore, in this paper, the viewpoint classification by using useful road features—comprising the vanishing point, the vertical line, the horizontal line, and the road area for different scenarios of bus waiting on the roadside—is shown.

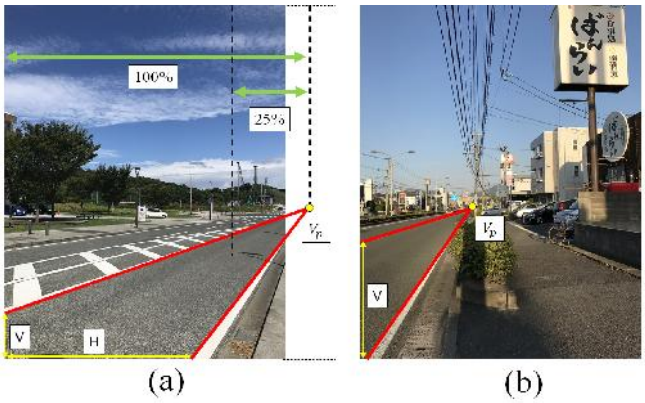

Fig. 5 (a) Example of suitable camera direction and its features, (b) Example of unsuitable camera direction and its features.

III. Proposed Method

The proposed method could discriminate image viewpoints when a blind individual takes a photo using his or her smartphone when waiting for a bus on the roadside. In fact, there are several road viewpoints that can appear when a user takes the photos; however, in this paper, only a few cars are assumed to be present on the road (non-congested traffic). This research assumes that the smartphone is held vertically using an inertial sensor data inside the smartphone. In addition, the viewpoints of the images will be distinguished into two main categories that consist of suitable and unsuitable image viewpoints for bus waiting.

The difficulty of outdoor scenes segmentation, especially road area, is uncertain illumination and unstructured shapes. Hence, this research presents the new combination method for road segmentation and useful features extraction for image viewpoint classification, as shown in Figure 7. Four main processes consist of finding the reference of road area, pre-viewpoint classification, road segmentation and feature extraction, and viewpoint classification.

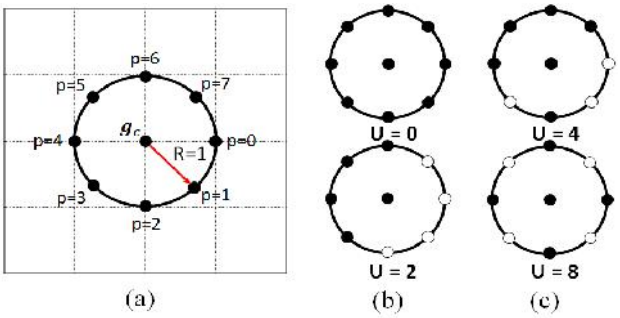

Fig. 6 (a) Example of $LBP^{riu2}_{8,1}$, (b) Example of Uniform LBP for U = 0 and U = 2, (c) Example of Ununiformed LBP for U = 4 and U = 8.

A. Finding the Reference of Road Area

Generally, roads have an unstructured shape and there are different illuminations for each image viewpoint. Hence, the first step in the proposed method is finding the reference of road area in the images. Image texture is one useful feature for image segmentation, especially the Rotational Invariant of Uniform Local Binary Pattern $LBP^{riu2}_{P,R}$, as proposed by T. Ojala *et. al.* [17]. $LBP^{riu2}_{P,R}$ is the technique for texture analysis of an images that finds the relation between the gray-scale value of the center (g_c) and its neighbors (g_p). In Figure 6a, the pattern of $LBP^{riu2}_{8,1}$ that this research defines, the number of neighbors (P) and radius (R) as 8 and 1 respectively, is shown. Each gray-scale value of neighbor (g_p) will be compared with g_c ; then, the value of 0 and 1 will be set for $g_p - g_c < 0$ and $g_p - g_c \geq 0$ respectively. Nevertheless, only uniform LBP ($U \leq 2$) is considered, where $U \leq 2$ is the pattern of LBP that has no transition (U = 0) or two transitions (U = 2), as shown in Figure 6b wherein the black and white points represent 0 and 1, respectively. For example, the binary code of 00000000 is defined by U = 0, and 10000111 is represented by U = 2. In case of non-uniformed pattern (U > 2), it is shown in Figure 6c that the binary codes are 00001011 and 10101010 for U = 4 and U = 8, respectively. According to uniform pattern limitation, there are P + 2 possible output values that can be represented at the center coordinate point.

This paper applies the $LBP^{riu2}_{8,1}$ method on the gray-scale image, as shown in Figure 8a. In Figure 8b, the complete transformation image of $LBP^{riu2}_{8,1}$, where each pixel is represented by a value between 0 and 9, is shown. Next, a sub-window of 30 × 30 pixel size is designed in order to calculate the histogram of $LBP^{riu2}_{8,1}$ transformation image for each sliding sub-window. Subsequently, each histogram of the sub-window will be grouped by using the k-means clustering method [18] in which k = 5 and Euclidean distance measurement, as defined in this research. Five different labels are separated and shown in Figure 9 with the representative histogram of each label (k_n). However, there is only one label that can be selected, and that label must contain the road area. Feedforward backpropagation of Artificial Neural Network (ANN) is applied for this selection. There are three main layers of ANN, which consists of input layer, hidden

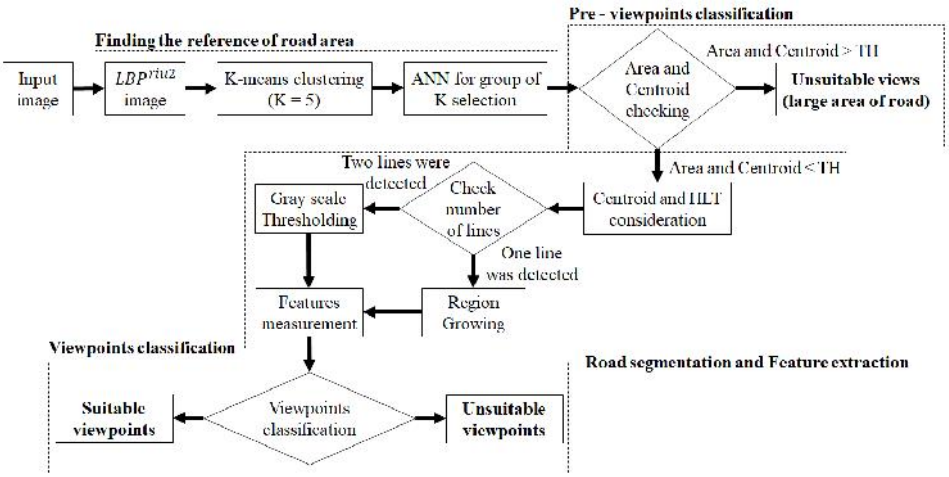

Fig. 7 Block diagram of the proposed method.

layer, and output layer. For the ANN design, the input layer comprises ten nodes for ten bins of histogram. Then, each node of the input layer is multiplied by different weight values and summarized in the hidden layer. Subsequently, the log-sigmoid transfer function is used for obtaining the value between 0 and 1. The output layer defines only one node for arriving at 0 or 1 that indicates whether the label does or does not contain the road area, respectively. For example, representative histogram of k_1 contains the road area; thus, the output k_1 is shown as one, whereas other representative histograms provide an output of zero. After, the group pixel on the left side and the bottom of the image is selected as the *reference road area* because the image viewpoints for bus waiting always appear in the road area in that part of the images.

Table 1 Performance testing for selecting the group of k-means by using ANN classification

Conditions	Accuracy (%)
Sunny	100
Cloudy	100
Before sunset / Sunrise	100
During raining in daytime	100
After raining in daytime	100
Nighttime	97
During raining in nighttime	97
After raining in nighttime	100
Non-road	96
Average of accuracy	98.89

B. Performance Testing of K-means Selection using ANN Classification

As the purpose of this research is to use a smartphone to find the suitable viewpoints for an oncoming bus, various conditions of the road that depend on different illuminations from daytime until nighttime can appear on an image. In order to ensure the performance of the proposed technique for selecting the group of k-means, the road area is examined in various illumination conditions. Eight conditions were tested, namely sunny, cloudy, before sunset (and sunrise), during

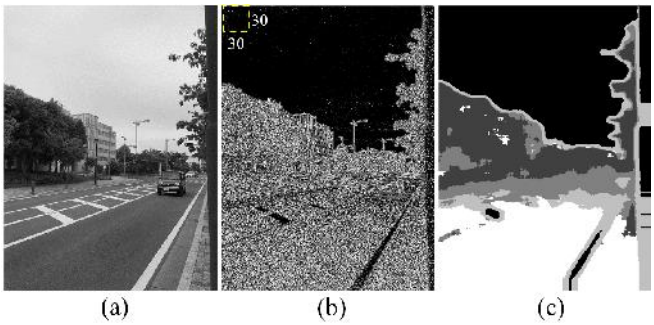

Fig. 8 (a) Gray-scale image with sub-window, (b) $LBP_{8,1}^{riu2}$ transformation image, (c) Image transformation from k-means clustering.

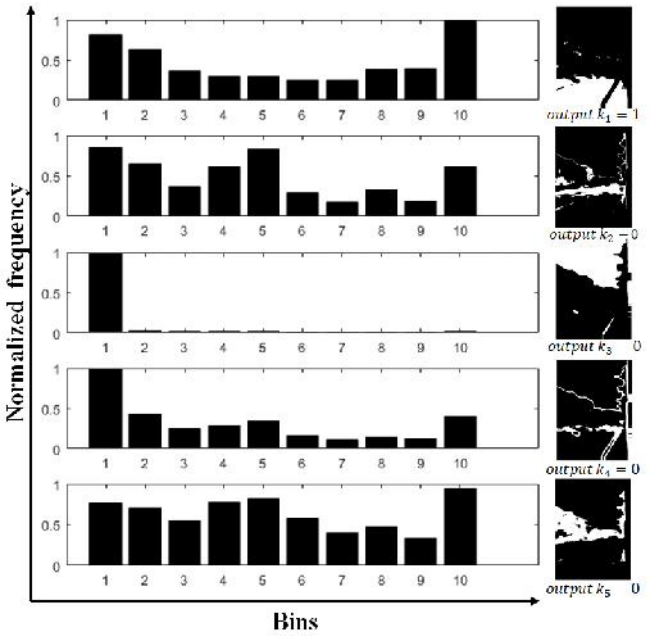

Fig. 9 Five different $LBP_{8,1}^{riu2}$ histogram representations from k-means.

raining in daytime, after raining in daytime, nighttime, during raining in nighttime, and after raining in nighttime. Likewise, the non-road areas were added for this testing in order to show the performance when ignoring non-road areas. For ANN training process, 400 different histograms for road and

non-road areas were used. In addition, 1,800 different histograms were tested for eight conditions of road and non-road areas. In Table 1, the percentage of accuracy for nine different conditions is shown; the average accuracy is 98.89%; thus, this percentage shows high effectiveness of the proposed method to select the road area.

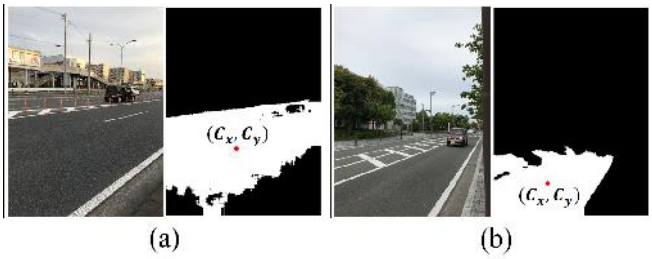

Fig. 10 (a) Example of unsuitable viewpoint for a very large road area case and its reference road area, (b) Suitable viewpoint and its reference road area.

C. Pre-Viewpoint Classification

In this section, the viewpoint classification for unsuitable viewpoint, in case of a large area of road, will be explained. In Figure 10a, one of the unsuitable viewpoints for bus recognition can appear when the area of the road is large. In Figure 10a and Figure 10b, the difference in binary images that is the reference road area is shown. There are two important features for separating the large area of road viewpoint from other views. The first feature is the percentage of reference of the road area because the large area of the road usually takes up more of the frame. Moreover, centroid coordinate point (c_x , c_y), as the red point in Figure 10, is calculated for the second feature.

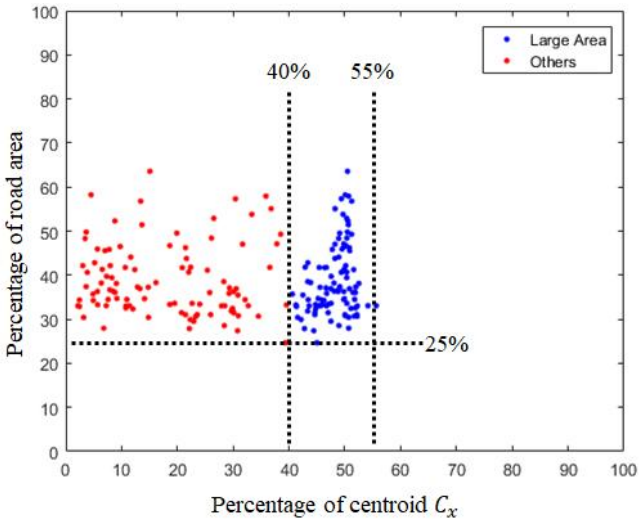

Fig. 11 Graph comparing features between large road area viewpoints and other viewpoints.

In order to set the criteria for large road area viewpoints, this research extracts two features of percentage of road area and percentage of road centroid c_x . A set of 100 sample images of large road area viewpoints and other views are used for feature consideration. In Figure 11, criteria setting for large road area—where the blue dots are the representative features of the percentage of road area and centroid c_x —is shown, clearly separating into two cases. Consequently, this research sets the criteria of large road area (LRA) case as shown in (1). However, the accuracy of pre-viewpoint classification will be shown in Section 4 of the experiments.

$$\text{case} = \begin{cases} \text{LRA}, \{(x, y) \mid 40 \le x \le 55 \text{ and } y \ge 25\} \\ \text{other}, \text{otherwise} \end{cases} \quad (1)$$

D. Road Segmentation and Feature Extraction

Until now, the LRA is classified from other cases that consist of suitable and unsuitable viewpoints. Thus, the next process comprises the method for road segmentation and feature extraction. Although the reference area of road can be found using the previous step, those segmented areas appear as a rough road area. Therefore, this research presents the technique that can be used to segment the road area clearly. Then, the necessary features of image viewpoint for roadside bus waiting will be extracted. For the road segmentation process, the two main steps comprise road edge detection using Hough Line Transformation (HLT) and line selection, followed by gray-scale and seed region growing segmentation.

(1) Road Edge Detection Using Hough Line Transformation and Line Selection

To border the road area, HLT is applied for detecting the straight-line markers on the road [19–21]. The HLT can transform each point in the (x, y) space to Hough space (θ, r). Although, θ can be obtained from −90 to 90 degrees for HLT, those are considered just 20 to 90 degrees because the possible marker lines on the image usually appear with those θ ranges. Figure 12a shows the image edge detection, which uses Sobel edge detection, which is a necessary step for the HLT transformation technique. The HLT technique is then used to detect the feasible lines, as shown in Figure 12b.

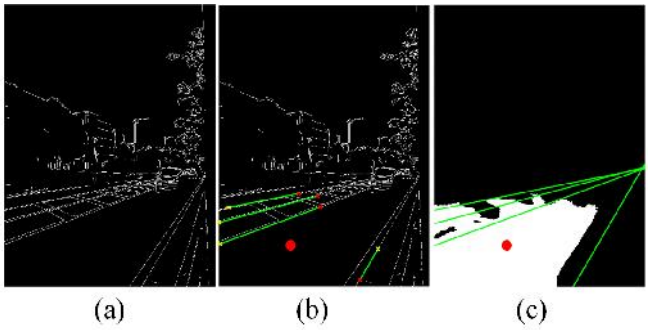

Fig. 12 (a) Sobel edge detection, (b) Line detection by HTL, (c) Reference road area and feasible lines.

Many possible lines appear in the image, as shown in Figure 12c. This study proposes the two border lines selection. Similar to the previous process, the initial road has been detected and the centroid of the road pixel area is calculated as the red point in Figure 12c. The shortest distance between centroid and those lines need to be determined. First, it can be assumed that the (h, k) in Figure 13(a) is the centroid of road area. In addition, L_1 represents the straight-line that has been detected by HLT previously. The point (x, y), which is the touching point between L_1 and circle edge with perpendicular line of L_2 , which can be found using (2). Equation (2), m_1 and m_2 are the slopes for L_1 and L_2 , respectively. Further, the distance, (r), between the centroid and (x, y) coordinate points can be calculated using (3).

$$m_1 \times m_2 = -1 \quad (2)$$

$$r = \sqrt{(x-h)^2 + (y-k)^2} \quad (3)$$

Each line is calculated using the proposed method, as shown in Figure 13(b) and Figure 13(c). The two lines are required, so the graph is separated into two areas comprising an upper and a lower centroid, as shown in Figure 13b. The minimum of r value for each area is selected. Furthermore, there are some cases where only one line of the road marker is detected because the marker line on the other side of the road might not be clear or might be obscured by cars. However, a similar method can be applied to find the shortest line of r for upper centroid, as shown in Figure 13c.

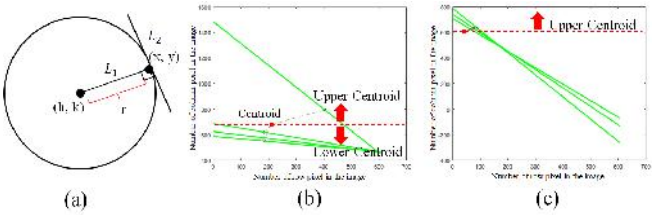

Fig. 13 (a) Line touching circle at x, y point, (b) Upper and Lower centroid separation for two-line detection, (c) Upper centroid selection for one-line detection.

(2) Gray-Scale and Seed Region Growing Segmentation

After defining the road area by detecting two lines, the next step is road segmentation. Gray-scale thresholding for outdoor segmentation is difficult to use, so this proposed method can be applied because the reference road area is roughly known. In Figure 14, the different histograms of gray-scale values that depend on illuminations of the reference road area are shown from the previous process. Nevertheless, some reference road areas might contain some small noises such as a small part of a white marker on the road or some small non-road area; these small noises usually appear as the small peak, as shown in Figure 14 (top right). In order remove these noise areas, this research sets the cut-off value at 90% differing with the highest peak for both sides, as shown as an example in Figure 14 (lower row). Subsequently, the remaining range of the histogram will be applied as the threshold value.

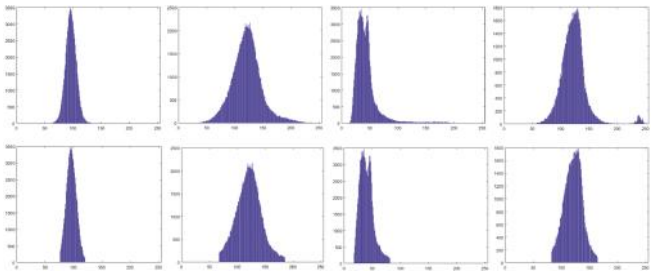

Fig. 14 Example of gray-scale histogram from different reference road areas (upper row) and cut-off the histogram range by 90% differing with the highest peak for both side (lower row).

There are two cases of line detection, namely two-line detection and one-line detection. In the case of two-line detection, the gray-scale thresholding will be used for road segmentation within the enclosed area by selecting two lines, as shown in Figure 15a. On the other hand, the seed region growing technique [22–23] is applied to segment the road area in the case of one-line detection, as shown in Figure 15b. The initial seed is the centroid of the reference road area. Each seed will extend to the four nearby directions with the threshold value mentioned in the previous proposed technique. The result of applying this technique is shown in Figure 15d.

Table 2 List of features extracted with their significance

Cases of lines detection	Features	Trend of value	
		Suitable viewpoints	Unsuitable viewpoints
1 and 2	% of V / % of H	Very small	Very large
2	% of Vp for x-coordinate	Large	Small
2	% of Vp for y-coordinate	Large	Small
1 and 2	% of road area	Large	Small

(3) Feature Extraction and Viewpoint Classification

Because the purpose of this research is to ensure suitable viewpoint classification of bus waiting for blind individuals, this paper proposes the necessary road features that consist of four features. In Figure 15c, three road features that comprise the vertical length of road (V), horizontal length of road (H), and the road areas shown. Moreover, vanishing point (Vp), as shown in Figure 15a, is added for the case of two-line detection.

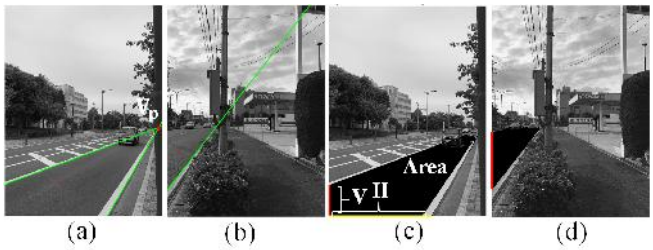

Fig. 15 (a) Detection of two lines and vanishing point (V_p), (b) Detection of one line, (c) Three road-feature components by gray-scale thresholding segmentation, (d) Road area segmentation using Seed region growing technique.

Table 3 Accuracy of pre-viewpoint classification using criteria setting

Viewpoint types	Accuracy (%)
Large road area	98
Other	100

The definitions of suitable and unsuitable viewpoints discussed in Section 2 are extracted and calculated in terms of the ratio of percentage of vertical length (V) and percentage of horizontal length (H), the percentage of vanishing point for the x coordinate, the percentage of vanishing point for the y coordinate, and the percentage of road area. Table 2 shows the list of features extracted with their significance of trend of value between suitable and unsuitable viewpoints. The ratio of percentage of V and H is usually very small and large for suitable and unsuitable viewpoints, respectively. The percentage of Vp for x and y coordinates of suitable viewpoints are larger than the unsuitable viewpoints. In addition, the percentage of road area usually shows the small

Table 4 Performance testing for feature extraction

Conditions	% Error, S.D.					Average of % error, S.D. (Conditions)
	Vertical	Horizontal	Vp(x)	Vp(y)	Area	
Sunny	6.57, 10.17	11.40, 18.17	1.90, 4.02	1.05, 1.61	7.73, 10.44	5.73, 8.88
Cloudy	3.73, 8.48	10.56, 19.88	1.06, 1.02	0.92, 1.10	6.04, 8.38	4.46, 7.77
Before sunset/ Sunrise	2.81, 4.48	3.30, 5.69	1.35, 2.88	1.33, 1.83	3.33, 5.44	2.42, 4.06
During raining (daytime)	2.05, 4.75	5.94, 15.96	2.91, 6.23	1.91, 3.03	3.08, 4.30	3.17, 6.85
After raining (daytime)	0.81, 0.67	0.85, 0.97	0.96, 1.20	1.12, 2.01	1.06, 0.83	0.96, 1.14
Nighttime	1.40, 1.12	9.29, 17.23	4.08, 12.28	2.20, 4.03	5.78, 5.77	4.55, 8.09
During raining (nighttime)	5.25, 16.97	8.25, 10.13	4.78, 8.40	4.51, 5.97	8.99, 11.89	6.36, 10.67
After raining (nighttime)	1.28, 1.07	3.52, 7.43	0.86, 1.47	0.81, 0.78	2.66, 2.64	1.83, 2.68
Average of % error, S.D. (Features)	2.99, 5.96	6.64, 11.94	2.24, 4.68	1.73, 2.56	4.83, 6.21	

value for unsuitable viewpoints but a large value for suitable viewpoints. However, the Vp for x and y coordinates in case detection of one line will be set as zero value because it cannot find the vanishing point.

For the classification process, the ANN of feedforward backpropagation is applied in which the input layer, the hidden layer and the output layer of ANN are designed by four, ten, and one nodes respectively. The four nodes of the input layer consist of the four feature values as shown in Table 2. Furthermore, the output layer is designed with one node that can return the value of one or zero for suitable viewpoints or unsuitable viewpoints, respectively. However, the performance of feature extraction and viewpoint classification with proposed method will be shown in Section 4 of the experiments.

IV. Experiments

The performance of the proposed method will be included in this section. Three different experiments consisting of pre-viewpoint classification, feature extracted performance, and viewpoint classification were carried out. Though, our proposed method aims to solve the problem for Thailand, this research collected the images from scenarios of roadside viewpoints under various illumination conditions in Japan between day and night as shown in Figure 16. Because of the same road type between Thailand and Japan, thus we can assume and collect the sample images for this experiment. In addition, images were captured by a smartphone with the smartphone being held vertically. The input image size was 800 px × 600 px, RGB color format.

Fig. 16 Example of roadside viewpoints of different illumination conditions.

A. Pre-Viewpoint Classification Accuracy

Applying the proposed method, using the criteria setting of the percentage of road area and the c_x coordinate of the vanishing point, two hundred images for the large area of road viewpoints and other viewpoints, as shown in Figure 10, were collected. Table 3 shows the experimental results of accuracy testing between LRA viewpoints and other viewpoints. The accuracy percentages for LRAs and other viewpoints are 98 and 100, respectively. There are two percent errors for LRA viewpoints because a certain percentage of road area and the c_x coordinate of vanishing point are outside of the criteria setting. However, the accuracy is still high with the proposed criteria setting.

B. Feature Extraction Performance

For the experiment of feature extraction, eight different conditions, as shown in Table 1, were tested. Five features were extracted, that is, vertical length of road, horizontal length of road, vanishing coordinate point (Vp(x), Vp(y)) and area of road. Four hundred images were tested for all conditions, and five features were extracted for each image. Then, the extracted feature values were compared with actual values that were provided by counting manually. The error percentage was calculated by below equation.

$$\%\text{Error} = \left| \frac{\text{actual value} - \textit{exp}\text{eriment value}}{\text{actual value}} \right| \times 100 \quad (4)$$

Table 4 shows error percentage and standard deviation (S.D.) for each condition and feature. All errors were lower than 7.00%, in both terms of conditions and features measurement. In term of conditions, the biggest error and S.D. are 6.36% and 10.67, which is when it is raining at night time. The roadside lights and vehicle headlights during rain can affect the unstable illumination of road area. Similar to sunny conditions, it shows the second highest error by 5.73% because this condition can appear various patterns of illumination from sunshine.

On the other hand, the condition after rain (daytime) shows the lowest error value by 0.96% and 1.14 of the S.D. value. This is because the road area after daytime rains is quite homogeneous, owing to the road being wet without sunshine. Therefore, the performance of feature extraction is high in post rain conditions during daytime.

In terms of features, five extracted features will be applied for viewpoint classification. The highest error is 6.64% for horizontal length feature, but the feature of vanishing

coordinate point (Vp(y)) shows the lowest error at 1.73% and vanishing coordinate point (Vp(x)) shows the second lowest error at 2.24%. This means that the performance of the vanishing coordinate point (Vp(x), Vp(y)) extraction method is highly accurate compared with the other four features of extraction. Both error measurements for conditions terms and features terms show high performance, as seen in Table 4. However, in Section 5 of the discussion, the results for viewpoint classification will be confirmed.

C. Viewpoint Classification Accuracy

As the purpose of this paper is to find the suitable viewpoints of bus waiting by the roadside, the quantitative measurement is necessary. For this experiment section, all images were collected from the different bus stops around Wakamatsu-ku, Kitakyushu city, in Japan. The number of 280 images were provided for training the feedforward backpropagation of ANN. These training images include both suitable and unsuitable viewpoints that were considered as proposed techniques in Section 2. After, all features were extracted by the previous method and data was sent to the ANN for training process.

For the testing process, 800 images as unknown data were prepared for 8 different conditions. In addition, each condition was half separated for suitable and unsuitable viewpoints. Then, ANN as previously mentioned, was used to classify all unknown data of images. The output of ANN showed the value (one or zero), and one or zero meant suitable or unsuitable viewpoints for bus waiting, respectively. Subsequently, the output from the ANN classification was compared with suitable definition, as explained in Section 2. In Table 5, the quantitative results for final performance of viewpoint classification is presented.

Table 5 Accuracy of viewpoint classification

Conditions	Accuracy of viewpoint classification (%)		Average of % accuracy
	Suitable	Unsuitable	
Sunny	96	100	98
Cloudy	100	100	100
Before sunset/ Sunrise	100	100	100
During raining (daytime)	98	100	99
After raining (daytime)	100	100	100
Nighttime	96	100	98
During raining (nighttime)	96	100	98
After raining (nighttime)	99	100	99
Total average of accuracy	98.13	100.00	99.00
Standard deviation	1.89	0.00	0.93

According to the experimental results in Table 5, the highest accuracy of 100% in three conditions, namely cloudy, before sunset, and after raining in daytime is shown. Nevertheless, the lowest accuracy was seen in three conditions of sunny, nighttime, and during rains in the nighttime—at 98%. Further, the unsuitable viewpoints can be classified perfectly by 100% in all conditions. However, some mistakes in suitable viewpoint classification are seen whereby the total average accuracy is only 98.13%, as shown in Table 5.

V. Discussion And Future Work

A. Discussion

This paper proposed the novelty of viewpoint classification for non-congested traffic. The advantage of this proposed viewpoint classification is to complement existing research that solely focused on bus detection and bus number recognition. Moreover, this paper applies the combined simple techniques of image processing and deep learning. Our proposed application will apply to smartphone with real-time process, therefore, the simple technique with light load calculation is feasible and suitable. Although, a simple technique of image processing is used, all experimental results showed high performance in Section 4.

However, in this section, the differences between error results of feature extraction process (Table 4) and accuracy of viewpoint classification (Table 5) will be described as shown in Figure 17. The error results for each condition from Table 4 are represented in Figure 17a. The condition after raining (day) has the lowest error compared to others. Although an error in the feature extraction process appears in the graph, as shown in Figure 17a, the accuracy of viewpoint classification is still high in all conditions, as shown in Figure 17b.

As the lowest average of error value (0.96%) corresponds to the condition after rain in daytime, the highest accuracy for this condition is seen at 100%. In contrast, the lowest accuracy of viewpoint classification can be seen in three conditions of sunny, nighttime, and during rain in the night—at 98%, which is the biggest error of feature extraction corresponding to rain in the nighttime at 6.36%. Furthermore, 5.73% and 4.55% are the second and third highest errors for sunny skies and raining at night conditions, respectively.

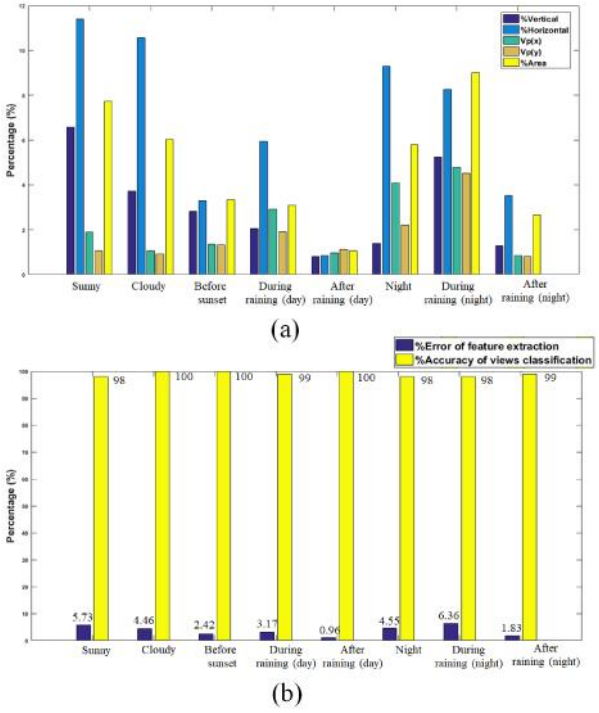

Fig. 17 (a) Graphical representation of error percentage for feature extraction, (b) Compared graphical between viewpoint classification and error of feature extraction.

B. Future Work

Although, these results show high performance, there are some mistakes in the viewpoint classification process. For suitable viewpoint classification, especially, wrong results can come about when comparing the proposed definition of suitable viewpoints for bus waiting described in Section 2. For example, in Figure 18, an example error case of viewpoint classification is shown. According to definition, the suitable viewpoints for bus waiting are shown in Figure 18a. However, this classification of viewpoints is inaccurate

because of wrong line selection, as shown in Figure 18b. The process of HLT can be detected, so many lines are shown in Figure 18c, but can be mistaken for some cases of the line selected process. However, this problem might be solved in future work by using some pre-processing technique for reducing the unrelated lines. Moreover, the process of obstacle detection and the case of viewpoints when many cars are on the road will be considered in future work.

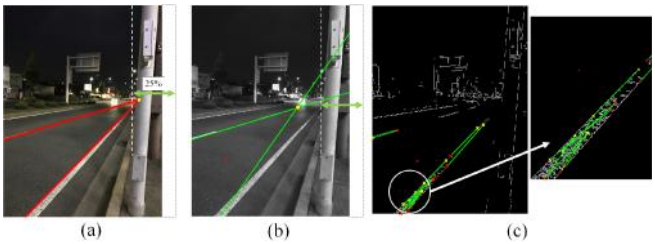

Fig. 18 (a) Good viewpoints image for bus waiting as defined in Section 2, (b) Example of mistake for line selection, (c) Many detected lines of HLT process.

VI. Conclusion

This paper proposed the novel application of viewpoint classification in cases of non-congested traffic conditions related to bus waiting, an idea that can be of great help to blind individuals in Thailand. The road area is detected for extracting the necessary features using a combination of methods. This technique is robust across various illuminations. There are four main processes that constitute the proposed method, namely finding the reference of road area, pre-viewpoint classification, road segmentation and feature extraction, and viewpoint classification. Based on this method, real scenarios of roadside viewpoints for bus waiting with eight different conditions in daytime and nighttime were tested. The experimental results show the high accuracy of viewpoint classification by 99% and 0.93 for S.D. Finally, we need to consider the case of congested traffic for viewpoint classification and how this proposed idea will be implemented in real-time applications on smartphones in the future.

References

[1] Ministry of Social Development and Human Security (2017). https://www.m-society.go.th/ewt_news.php?nid=20586. Accessed 4 Sept 2017.
[2] Thailand Association of The Blind (2016). http://tabgroup.tab.or.th/node/124. Accessed 16 Feb 2016.
[3] Dona S, (2008) Accessibility and Traffic Control in Asia. http://www.sauerburger.org/dona/photoaccess.htm. Accessed 1 Apr 2008.
[4] Chen-Fu L (2016) Accessible Traffic Signals for Blind and Visually Impaired Pedestrians. http://www.me.umn.edu/~cliao/PDF/Mobile_APS_OTC.pdf. Accessed 1 Mar 2016.
[5] Santos EAB (2015) Design of an interactive system for city bus transport and visually impaired people using wireless communication, smartphone and embedded system. In: 2015 Proceeding of the SBMO/IEEE MTT-S Microwave and Optoelectronics Conference (IMOC), pp 1–5.
[6] El Alamy L, Lhaddad S, Maalal S, Taybi Y, and Salih-Alj Y (2012) Bus identification system for visually impaired person. In: 2012 Proceeding of the IEEE International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST). pp 13–17.
[7] Venard O, Vaudoin G and Uzan G (2009) Field experimentation of the RAMPE interactive auditive information system for the mobility of blind people in public transport: Final evaluation. In: 2009 Proceeding of the IEEE International Conference on Intelligent Transport Systems Telecommunications, (ITST). pp 558–563.
[8] Noor MZH, Ismail I and Saaid MF (2009) Bus Detection Device for the Blind Using RFID Application. In 2009 Proceeding of IEEE International Colloquium on Signal Processing and Its Applications. pp 247–249.
[9] Kalbani JA, Suwailam RB, Yafai AA, Abri DA and Awadalla M (2015) Bus detection system for blind people using RFID. Proceeding of the IEEE International Conference on GCC Conference and Exhibition. p 1-6.
[10] Zhou P, Zheng Y and Li M (2014) Zhou, Pengfei, Yuanqing Zheng, and Mo Li. "How long to wait? Predicting bus arrival time with mobile phone based participatory sensing. IEEE Trans. Mobile Computing. 13(6):1228–1241.
[11] Digital Age Magazine [Thai] (2017) Introduction of ViaBus Application. ISSUE 216 in December.
[12] Wongta P, Kobchaisawat T, and Chalidabhongse TH (2016) An automatic bus route number recognition. In: 2016 in Proceeding of the IEEE International Conference on Computer Science and Software Engineering (JCSSE). pp. 1–6.
[13] Cheng CC, Tsai CM, and Yeh ZM (2014) Detection of bus route number via motion and YCbCr features. In: 2014 Proceeding of the IEEE International Conference on Computer, Consumer and Control (IS3C). pp 31–34.
[14] Lee D, Yoon H, Park C, Kim J and Park CH (2013) Automatic number recognition for bus route information aid for the visually-impaired. In: 2013 Proceeding of the IEEE International Conference on Ubiquitous Robots and Ambient Intelligence (URAI). pp 280–284.
[15] Guida C, Comanducci D and Colombo C (2011) Automatic bus line number localization and recognition on mobile phones—a computer vision aid for the visually impaired. In: 2011 Proceeding of the International Conference on Image Analysis and Processing. pp 323–332.
[16] Pan H, Yi C, and Tian Y (2013) A primary traveling assistant system of bus detection and recognition for visually impaired people. In: 2013 Proceeding of the IEEE International Conference on Multimedia and Expo Workshops (ICMEW). pp 1–6.
[17] Ojala T, Pietikainen M and Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell 24(7):971–987.
[18] Arthur D and Vassilvitskii S (2007) k-means++: The advantages of careful seeding. In: 2007 Proceeding of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. pp 1027–1035.
[19] Hough, Paul VC. "Method and means for recognizing complex patterns." U.S. Patent No. 3,069,654. 18 Dec. 1962.
[20] Duda RO and Hart PE (1972) Use of the Hough transformation to detect lines and curves in pictures, Commun. ACM. 15(1): 11–15.
[21] Ballard DH (1981) Generalizing the Hough transform to detect arbitrary shapes. Pattern Recognit. 13(2):111–122.
[22] Adams R and Bischof L (1994) Seeded region growing. IEEE Trans. Pattern Anal. Mach. Intell. 16(6):641–647.
[23] Fan J, Zeng G, Body M, and Hacid MS (2005) Seeded region growing: an extensive and comparative study. Pattern Recognit. Lett., 26(8):1139–1156.

Watcharin Tangsuksant received his B.Eng. degree in Biomedical Engineering from Srinakharinwirot University, Bangkok, Thailand, in 2013 and his M.Eng. degree in Biomedical Engineering from King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand, in 2015. He was a Lecturer with Rangsit University, Pathum Thani, Thailand, in 2016. He is currently a Ph.D. student of the Graduate School of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research interests include the image processing, signal processing, and assistive technology for disabled people.

Chikamune Wada Wada received the B.Eng. degree in mechanical engineering from the Osaka University, Japan, in 1990 and the Ph.D. degree in biomedical engineering from Hokkaido University, Japan, in 1996. From 1996 to 2001, he was an Assistant Professor with the Sensory Information Laboratory, in Hokkaido University. In 2001, he became an Associate Professor with Human-function Substitution System Laboratory, Kyushu Institute of Technology. Since 2016, he has been a Professor with Human-function Substitution System Laboratory. His research interests include assistive technology, especially measuring human motion and informing the disabled people of the necessary information to improve their QOLs. He is a senior member of the Institute of Electronics, Information and Communication Engineers (IEICE).