

Road Traffic Signs Recognition Using Genetic Algorithms and Neural Networks

Stephen Karungaru, Hitoshi Nakano, and Minoru Fukumi

Abstract—This paper proposes an efficient real time driving sign “stop” detection method using template matching, genetic algorithms and neural network. Stop signs are usually installed on junctions without traffic lights to warn drivers. However, many accidents still occur at these locations because either the driver does not notice the signs or just ignore them. Therefore, to reduce such accidents, we propose an automatic stop sign detection method to aid the drivers and also contribute to future automatic driving system such as the Intelligent Traffic System (ITS). Our method efficiently extracts the sign region’s candidate regions, performs template matching using genetic algorithms and verifies the presence of the road sign using neural networks. Although, we face various problems including camera shake and complicated scenes, our method produces an encouraging accuracy of about 96%.

Index Terms—Genetic algorithms, road signs, template matching, neural networks.

I. INTRODUCTION

In recent years, although the number of deaths caused by traffic accidents has decreased, the number of accidents is on the rise. According to traffic authorities in Japan, based on the type of accident, rear-end collisions and collisions at intersection accounted for about 60 percent of the total in 2009. Some of the reasons attributed to this were mostly due to driver inattentiveness and reckless driving (especially approaching the junction fast). Consequently, looking at traffic accidents by road section, road intersections had the highest number. Moreover, the majority of the accidents occur at junctions without traffic lights. However, in most of such junctions, the stop sign is usually installed. Nonetheless, traffic accidents such as reckless driving by young people and crashing into the intersection by the elderly, caused by the failure to temporary stop never decrease. Therefore, if we can automatically detect road signs, it can prevent oversight and be expected to reduce traffic accidents and violations. For this reason, in this study, we propose a method to detect the stop sign using a dashboard camera and image processing.

Road sign detection is actively studied by many researchers [1]-[5]. Above all, many methods including template matching have been proposed. However, template matching possess low speeds because, it requires round-robin comparison of template and input image pixels. It also may require a number of template images for various illumination conditions and sizes. Therefore, in this study, we use genetic algorithm to optimize the matching process and color

extraction to narrow the search range. Additionally, neural networks are used to confirm the results of the template matching process thereby improving the detection accuracy of our method.

The rest of this paper is organized as follows. The next section describes the concrete methods of stop sign detection and in Section III we describe the experiment and the result. Finally, Section IV describes the summary.

II. PROPOSED METHOD

Fig. 1 shows the overall block diagram of the flow of processing. At first, we input a road sign image that was taken from a car’s dashboard and we extract red color regions from the image.

We then remove noise from the resulting image using the median filter and labeling. Next, we detect an area that is similar to the stop sign using template matching and genetic algorithms. Finally, we judge whether the detected area is a stop sign using a neural network. The details of each process are described below.

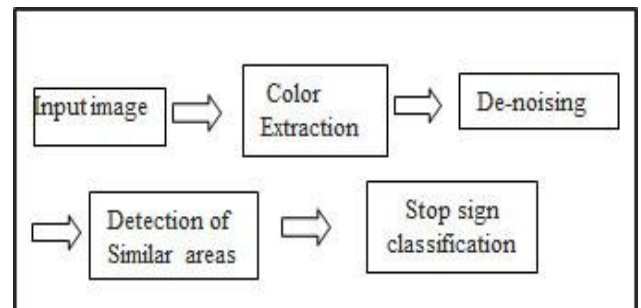


Fig. 1. Flow of the proposed method.

A. Color Extraction

In Japan, the stop sign is an inverted triangle that is mainly red in color. This color doesn’t blend in with most driving environments and therefore should be easy to spot. The red color is extracted using the following method.

In the most commonly used RGB model, it is difficult to understand the relations between luminance, hue and saturation. Consequently, we assume that it would be difficult to extract the red color using the RGB color model. Therefore, at first, we convert RGB color model to another color model. In this study, we use Lab color model because it has a high reproducibility of human vision. The following equations are used to convert from RGB to Lab color space.

$$X = 0.412453 \times R + 0.35758 \times G + 0.180423 \times B \quad (1)$$

$$Y = 0.212671 \times R + 0.71516 \times G + 0.072169 \times B \quad (2)$$

$$Z = 0.019334 \times R + 0.119193 \times G + 0.950227 \times B \quad (3)$$

Manuscript received May 18, 2012; revised April 15, 2013.

The authors are with the Graduate School of Advanced Technology and Science, The University of Tokushima (e-mail: karunga@is.tokushima-u.ac.jp).

$$L = 116 \times \sqrt[3]{x_r} - 16 \quad (4)$$

$$a = 500 \times \left(\sqrt[3]{x_r} - \sqrt[3]{y_r} \right) \quad (5)$$

$$b = 200 \times \left(\sqrt[3]{y_r} - \sqrt[3]{z_r} \right) \quad (6)$$

$$x_r = \frac{X}{0.95045}, y_r = Y, z_r = \frac{z}{1.08892} \quad (7)$$

X, Y, Z represents the tristimulus values.

“ L ” is a correlate of lightness. “ a ” roughly represents redness (positive) versus greenness (negative). “ b ” roughly represents blueness (positive) versus yellowness (negative).

Next, using four images that included stop signs, we made density histograms for “ L ”, “ a ” and “ b ”. Fig. 2 represents the density histogram of “ a ”. Fig. 2-(top) shows the entire image and Fig. 2-(bottom) only the stop sign area. Vertical axis represent the number of pixels and horizontal axis is the “ a ’s” value.

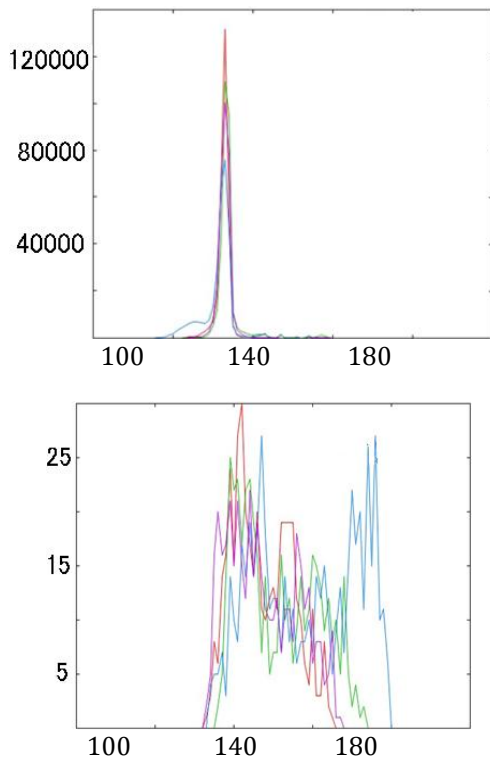


Fig. 2. Density histogram of “ a ”. Top: entire image; bottom: stop sign image.

These histograms show that many pixels have “ a ’s” value below 140 in the entire image (Fig. 2-top). On the other hand many pixels have “ a ’s” value over 140 in stop sign image. Therefore, we determine a preferable threshold to extract a stop sign from such images to be 140. Also, we determine the thresholds for “ L ” and “ b ” by the same histogram method. We shade areas below the thresholds black and those over it, candidate stop sign areas white.

B. De-Noising

The image extracted using the method above contains a lot of noise. We must remove such noise from the image. The first noise removal method we use is the median filter. The median filter can remove most of the fine noise. Next, we remove image components that have very small or large

width and length respectively using the labeling method in the median filter de-noised image.

Fig. 3 represents an example result of the noise reduction from image extracted using color information where Fig. 3-(top) is the original image. As shown, the search area is reduced to only a few black regions, Fig. 3-(bottom).



Fig. 3. Color extracted from the de-noised image. Top: Input image bottom: output image.

C. Detection of Similar Areas

After de-noising, template matching is performed within each of the remaining image components. Template images used are in nine different sizes as shown in Fig. 4.

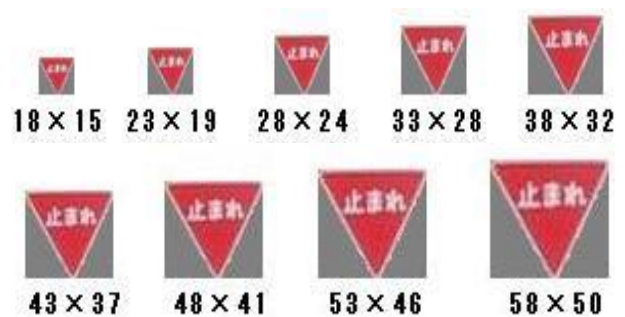


Fig. 4. Template images.

Within the search range, we calculate each component’s similarity with the template image. However, it takes a lot of time to process in all the coordinates of the search range. Therefore, to improve the speed, we use a genetic algorithm. Fig. 5 represents the flow of detection processing using genetic algorithm.

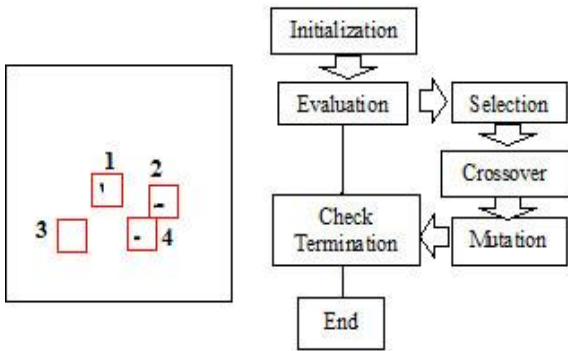


Fig. 5. Flow of detection process using genetic algorithm.

We work our way through areas 1 to 4 illustrated in Fig. 5 and calculate the maximum fitness in each of those areas. First, we process area 1.

The genetic algorithm [6] is set up as follows. We provide an individual carrying three genes. These are the X-Y coordinates and template image. Next, we initialize the chromosome to give a random X-Y coordinates and template image.

However, the X-Y coordinates are selected from area 1 and template image is selected from Fig. 3. Next, we calculate fitness of the individual using template matching. We perform template matching with a central focus on individual coordinates using the template image. We define the fitness as proportional to similarity that was calculated by template matching. Next, we perform crossover using the chromosomes that were selected according to fitness and finally mutate with a fixed probability. We repeat these steps and calculate the maximum fitness in area 1.

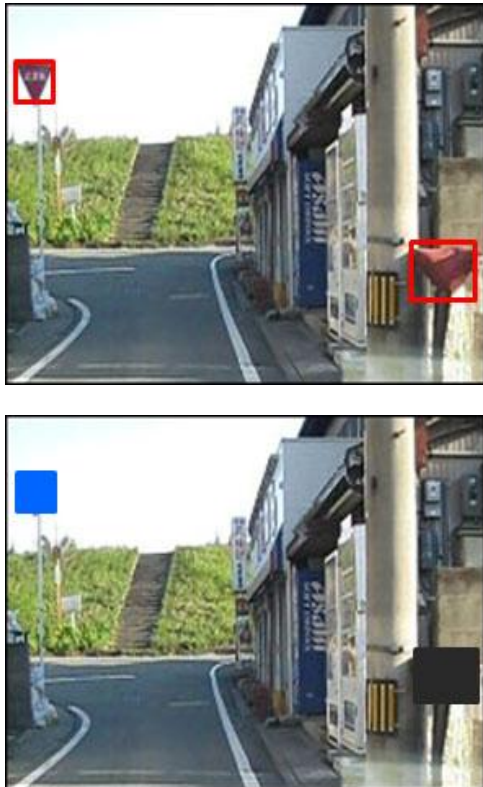


Fig. 6. Detection of similar areas. Top: input image bottom: output image.

The process is repeated in areas 2 to 4, calculating the maximum fitness in each of those areas. Finally, we compare

maximum fitness and save the top two elite individuals. We detect the area (size of template image) with a central focus on elite individual coordinates.

Fig. 6 shows an example result of the two detected areas that are similar to the stop sign. Fig. 6-(top) shows the two candidate areas extracted using color information. In Fig. 6-(bottom), the blue area is the area that was most similar to the stop sign. The black area had the second best result from the genetic algorithm and therefore is ignored because we assume one stop sign per image.

D. Stop Sign Classification

In this process, we use a neural network to verify whether or not the detected area is stop sign.

As a pre-processing step for this process, we resize the detected image to 20×20 pixels. The reason is so as to correspond to the number of neural network's input layer nodes, and reduce the number of dimensions. Next, we extract red color from the image that was resized. In the color extraction, we use the model proposed in Section II. The red pixels are extracted directly, while a white color is assigned to a pixel that wasn't red using threshold processing. Finally, the neural network inputs are "a" of the pre-processed image. Sigmoid function is used as an evaluation function.

III. EXPERIMENT

Using the proposed method, we performed experiments on 47 drive scenes that were taken on a camera placed on the dashboard of a car moving at between 30 and 40km/h. All these images taken include the stop sign. In this section, we describe the experiment, results and consideration.

A. Experiment 1

We performed the detection experiment using proposal method (Section II-A, B, C). If it detected a stop sign just once when the car is more than 13 meters away from the stop sign, we consider the detection successful. 13 meters is the average stopping distance when the car runs at 30 kilometers per hour.

B. Experiment Conditions

- Image size: 640x480pixels
- Frame rate:5 frame per second
- Number of drive scene: 47 (Each about 20secs long)
- Number of individuals used in genetic algorithm: 26
- Number of generations used in genetic algorithm: 20

C. Results of Experiment 1

Table I shows the result of experiment 1. From a total of 47 scenes, we successfully detected the stop sign in 45 scenes.

TABLE I: RESULTS OF EXPERIMENT 1

Total samples	Identify	Failed
47 scenes	45 scenes	2 scenes
Percentage	96%	4%

Moreover, after narrowing the range of searching using

color extraction, we also did the experiment with only template matching (not using genetic algorithm) for comparison. The result for both experiments was the same. But, there was a large difference in processing time.

Table II shows the processing time. The processing time using a single template matching is about 6.0 seconds. On the other hand, using genetic algorithm, the time reduces to only 0.3 seconds.

TABLE II: PROCESSING TIME

Method	Processing time
Using genetic algorithm (propose method)	About 0.3 seconds
Without genetic algorithm	About 6.0 seconds

D. Experiment 2

In this experiment, we used a neural network [7] to verify that the stop signs detected by the genetic algorithm are correct. We did the identification experiment using the 45 scenes that were successfully detected in experiment 1. The neural network chosen is a three layer structure (only one hidden layer).

We used half of these scenes for learning, and used the remaining half for evaluation. That is, divide 45 scenes into 23 and 22 scenes at random, and evaluate (learn) them alternately. Also, as a teacher signal, we represent the stop sign image using 0.96, and non-stop sign image using 0.04. We set the threshold of the output value to 0.8, that is, if the output is above 0.8, we take the result as a stop sign. Moreover, if the number of frames successfully identified is more than the number of frames that failed, we consider the detection successful.

E. Experiment Conditions

- Image size: 640 480
- Frame rate: 5 frame per second
- Number of drive scene: 45(22 and 23)

F. Results of Experiment 2

Table III shows the result of the experiment 2. From a total of 45 scenes, we successfully identified the stop sign in 43 scenes.

TABLE III: RESULTS OF EXPERIMENT 2

Total samples	Identify	Failed
45 scenes	43 scenes	2 scenes
Percentage	96%	4%

G. Consideration

Experiment 1: The experiment proved that the proposed method can shorten the processing time while maintaining the detection accuracy. Therefore, using genetic algorithm is considered of value in the detection process using template matching. Also, we consider the cause of the failure detection. Fig. 7 shows the image in which the detection failed. We couldn't detect the stop sign in this image at all because of the background sunlight. The sunlight made it difficult to segment the candidate regions. We think it is difficult to

detect the stop sign that is affected by environmental changes such as sunlight. Therefore, we need to prepare the proposed method to improve detection accuracy even under to such environmental changes.



Fig. 7. Failed to detect.

Experiment 2: We consider the cause of the failed recognition in experiment 2. Fig. 8 (a) shows an image that failed to correctly be identified. In the image, we find that the image quality is very poor. So, this is attributed to the fact that image quality is severely affected by resizing processing. Therefore, we need to think about the size of detected image.

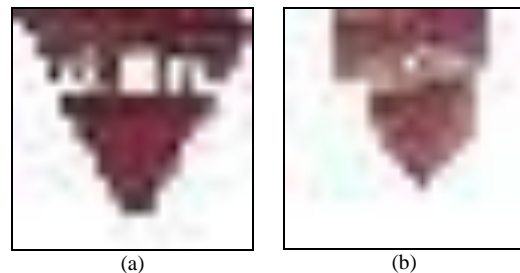


Fig. 8. Failed to accurately detect.

Also, the proposed method mis-recognizes the non-stop sign (such as Fig. 8-(b)) as the stop sign in many cases. But, we think it can be improved by increasing non-stop sign images in the learning data.

IV. CONCLUSION

In this study, we proposed a method to detect and identify the stop sign using image processing. If we can automatically detect road signs, it can prevent missed signs at junctions without traffic lights and this can be expected to reduce traffic accidents and traffic violations.

In the proposed method, first, we at first extract red color from the image. Secondly, we remove noise from the image using a median filter and labeling. Thirdly, we detect areas that are similar to the stop sign using template matching and genetic algorithm. Finally, we judge whether the detected area is a stop sign or not using neural network. The experimental result achieved 96% detection accuracy and 96% confirmation rates. Moreover, we found that using genetic algorithm can shorten the processing time while maintaining the detection accuracy. Therefore, the proposed method is considered useful.

The future challenge is to find better characteristics of stop

signs that can be used to make our method more robust to environmental changes.

REFERENCES

- [1] I. Arihito, F. Hironobu, T. Masanari, K. Hiroaki, and T. Yukimasa, "Improvement in the Accuracy of matching by different feature subspaces in traffic sign recognition," *IEEJ Trans. on Electronics, Information and Systems*, vol. 129, no. 5, pp. 893-900, 2009.
- [2] T. Masanari and F. Hironobu, "Traffic sign recognition using sift features," *IEEJ Trans. on Electronics, Information and Systems*, vol. 129, no. 5, pp. 824-831, 2009.
- [3] K. Doman, D. Deguchi, T. Takahashi, Y. Mekada, I. Ide, and H. Murase, "Cascaded traffic sign detector using generative learning considering color variance," *Technical report of IEICE. PRMU*, vol. 108, no. 363, pp. 135-140, 2008.
- [4] S. Miyata, A. Yanou, H. Nakamura, and S. Takehara, "Feature extraction and recognition for road sign using dynamic image processing," *Research reports of the Faculty of Engineering*, Kinki University, vol. 41, pp. 125-131, 2007.
- [5] I. Laptev and T. Lindeberg, "Space-time interest points," in *Proc. IEEE International Conf. on Computer Vision*, 2003, pp. 432-439.
- [6] A. T. F. M. A. Norio, "High Speed lips region extraction by template matching with genetic algorithms in real environments," *Technical report of IEICE. HIP*, vol. 104, no. 168, pp. 57-62, 1 July, 2004.
- [7] M. Ishikawa, "Structure learning with forgetting, neural networks," vol. 9, no. 3, pp. 509-521, 1993.



Stephen Karungaru received the Ph.D. degree in Information System Design from the Department of information science and Intelligent Systems, University of Tokushima in 2004. He is currently an associate professor in the same department. His research interests include pattern recognition, neural networks, evolutionary computation, image processing and robotics. He is a member of IACSIT, RISP, IEEE and IEEJ.



Hitoshi Nakano received a master degree from the Department of Information Science and Intelligent Systems, University of Tokushima in 2011.



Fukumi received the doctor degree from Kyoto University in 1996. Since 1987, he has been with the Department of Information Science and Intelligent Systems, University of Tokushima. In 2005, he became a professor in the same department. He received the best paper award from the SICE in 1995. His research interests include neural networks, evolutionary algorithms, image processing and human sensing. He is a member of the IEEE, SICE, IEEJ, IPSJ, RISP and IEICE.