

Improving the performance of traffic sign detection using blob tracking

Aryuanto Soetedjo^{1a)} and Koichi Yamada²

¹ Multimedia University,

Jalan Persiaran Multimedia 63100, Cyberjaya, Selangor, Malaysia

² Nagaoka University of Technology,

1603–1 Kamitomioka, Nagaoka, Niigata 940–2188, Japan

a) aryuanto@mmu.edu.my

Abstract: We propose a new approach for tracking circular traffic signs from image sequences to improve the performance of traffic sign detection, by reducing search region and suppressing misdetection caused by temporal occlusion or poor quality of image. Our proposed tracking, called *two-layered blobs tracking*, does not require an accurate model of the fixed object-moving camera system, which is essential in the Kalman-Filter tracking. The experimental results show that the proposed approach could track the circular traffic signs from a moving camera effectively, without any restrictions on speed and movement of the vehicle, and camera installation, thus it is easy to be implemented in real situation.

Keywords: traffic sign detection, traffic sign tracking, two-layer blobs, blob tracking, ring-partitioned matching

Classification: Science and engineering for electronics

References

- [1] Y. Aoyagi and T. Asakura, "A study on traffic sign recognition in scene image using genetic algorithms and neural networks," *Proc. 22nd IEEE IECON International Conference on Industrial Electronics, Control, and Instrumentation*, pp. 1838–1843, 1996.
- [2] A. de la Escalera, J. M. Armingol, and M. Mata, "Traffic sign recognition and analysis for intelligent vehicles," *Image and Vision Computing*, vol. 21, no. 3, pp. 247–258, March 2003.
- [3] C. Y. Fang, S. W. Chen, and C. S. Fuh, "Road-Sign Detection and Tracking," *IEEE Trans. Veh. Technol.*, vol. 52, no. 5, pp. 1329–1341, Sept. 2003.
- [4] G. Piccioli, E. De Micheli, P. Parodi, and M. Campani, "Robust Method for Road Sign Detection and Recognition," *Image and Vision Computing*, vol. 14, no. 3, pp. 209–223, April 1996.
- [5] J. Miura, T. Kanda, and Y. Shirai, "An Active System for Real-Time Traffic Sign Recognition," *Proc. IEEE Int. Conf. On Intelligent Transportation Systems*, pp. 52–57, 2000.
- [6] L. M. Fuentes and S. A. Velastin, "People Tracking in Surveillance Applications," *Image and Vision Computing*, vol. 24, no. 11, pp. 1165–1171,

- Nov. 2006.
- [7] L. A. Sechidis, P. Patias, and V. Tsioukas, “Low-Level Tracking of Multiple Objects,” *Proc. of ISPRS Commission III Symposium*, pp. 237–240, 2002.
 - [8] A. R. J. Francois, “Real-Time Multi-Resolution Blob Tracking,” *IRIS Technical Report*, IRIS-04-422, University of Southern California, Los Angeles, USA, April 2004.
 - [9] A. Soetedjo and K. Yamada, “Traffic Sign Classification Using Ring Partitioned Method,” *IEICE Trans. Fundamentals (Japanese Edition) Electronics*, vol. E88-A, no. 9, pp. 2419–2426, Sept. 2005.
 - [10] A. Soetedjo and K. Yamada, “An Efficient Algorithm for Traffic Sign Detection,” *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 10, no. 3, pp. 409–418, May 2006.

1 Introduction

Road traffic signs are designed to regulate traffics and provide useful information to drivers when they are driving on the street. Due to weather conditions, environment disturbances or human errors, misdetection of those signs could lead to traffic accidents. Traffic sign recognition systems, an intelligent vision system for recognizing traffic signs, will help the driver to recognize traffic signs early and accurately.

Many techniques have been developed to recognize traffic signs using a single image [1, 2]. Detection of traffic signs in only a single image has three problems [3]: a) information about positions and sizes of traffic signs, which is useful to reduce the computational time, is not available; b) it is difficult to detect a traffic sign correctly when temporary occlusion occurs; and c) correctness of the detection is hard to verify. In [3, 4, 5], image sequences were utilized for recognizing traffic signs. Tracking using the Kalman-Filters were adopted to track traffic signs in [3, 4] to allows a drastic reduction of the search region with respect to the single image case. In fact, if the location and size of a traffic sign in frame (t) is known, the search region in frame ($t + 1$) could be centered on the area around the predicted location and size of the sign as illustrated in Fig. 1.

Tracking in the previous works assumed that the vehicle moves straight with a constant speed [4, 5] or the speed of the vehicle and the physical sizes of road signs are known [3]. Violations of those assumptions are treated as errors of the supposed models. The Kalman-Filter works well when the accurate model is available. However, it is not common unfortunately in the real situation.

To overcome the difficulty of the approach with the Kalman-Filter, we propose a new approach called two-layered blobs tracking. Blob tracking is usually used to track moving objects like people, vehicles, or balls [6, 7, 8] by matching blobs defined as bounding boxes representing the foreground (moving) objects, from two successive frames. Our approach differs from the existing blob tracking [6, 7, 8] in two ways: a) our application tracks fixed

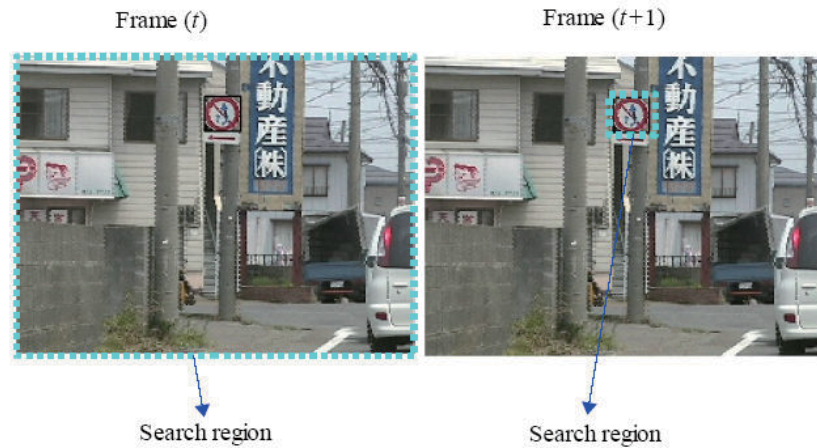


Fig. 1. Illustration of tracking for reducing search region.

objects from a moving camera, and b) our system could track blobs which are not in the successive frames.

2 Blob tracking

In the cases of the Kalman-Filters tracking, positions and sizes of traffic signs in frame $(t + 1)$ are predicted after the signs in frame (t) are recognized. Once they are predicted in frame $(t + 1)$, the search region to detect the signs is concentrated on the area determined based on those predictions. Our approach does not predict the positions and the sizes of signs in frame $(t + 1)$. Instead, blob tracking is used to assign blobs in frame $(t + 1)$ to the ones in the previous frame (t) or $(t - k)$, $k \geq 1$, which have already been recognized as traffic sign candidates. Using this method, traffic signs could be tracked without developing an accurate model for the prediction.

In the research, only circular signs with red color border and blue and white colors inside the circle are treated, however the proposed method could be expanded to cope with the other signs accordingly. At first, a red color thresholding [9, 10] is employed, and a connected component labeling is applied to create blobs layer-1. Using these blobs, the spatial information (bounding boxes) of signs is determined. Then a blue color thresholding [9, 10] is applied to sub-images inside the bounding boxes to create blobs layer-2. Blobs layer-1 and the blobs layer-2 are illustrated in Fig. 2.

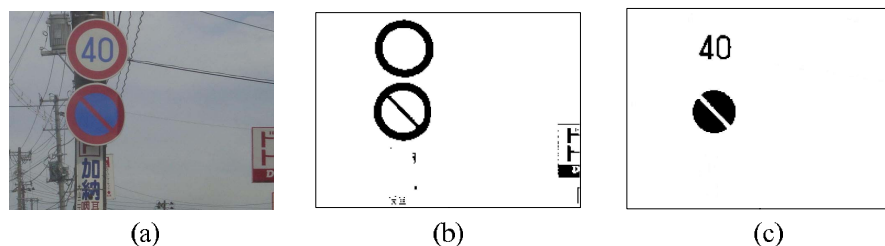


Fig. 2. Two layer blobs: (a) Original image; (b) Blobs layer-1; (c) Blobs layer-2.

Then, tracking is performed by matching blobs between frames. During tracking, the latest tracked blobs are stored in a temporary memory. Blobs in frame $(t + 1)$ are matched with those in the temporary memory, which holds the latest tracked blobs. Thus, when a sign disappears in particular frames, it could be tracked in the next frame when it appears again.

3 Blob matching

In the traffic sign tracking, signs are tracked from a camera that moves forward, possibly very fast. Therefore, the sizes of blobs are expected to increase from frame to frame, and the positions might change according to the movement of the vehicle. In addition, due to the temporal occlusion and/or missed sign detection, blobs in frame $(t + 1)$ may not have a corresponding blob in frame (t) , but may have the one in frame $(t - k)$, $k \geq 1$, in our system. With all of the above situations, the existing blob matching [6, 7, 8], which assume that the sizes and positions of blobs to be matched are close, would not work effectively.

Here, we propose blob matching using a ring-partitioned method. The ring-partitioned matching [9] provided a fast and robust matching of images under problems of scale varying, rotation, and partial occlusion. In the ring-partitioned matching, image is partitioned into several ring-shaped areas, and matching between the target and reference images is done by computing the distance or similarity of the normalized histogram of every ring of images. The distance between the blob in the current frame and the corresponding one is calculated as

$$d^{NO} = \sqrt{\sum_i w_i \left(\frac{pN_i}{pN_i + qN_i} - \frac{pO_i}{pO_i + qO_i} \right)^2} \quad (1)$$

where

pN_i is the number of black pixels of the current blob in the ring- i ;

qN_i is the number of white pixels of the current blob in the ring- i ;

pO_i is the number of black pixels of the corresponding blob in the ring- i ;

qO_i is the number of white pixels of the corresponding blob in the ring- i ;

i is a ring number ($i = 3$ in the experiment);

w_i is the weighting factor of the ring- i ($w_1 = 1$, $w_2 = 1$, $w_3 = 0.2$ in the experiment).

4 Experimental results

The proposed technique was evaluated to track red circular signs from image sequences obtained by a moving video camera installed on a vehicle. In the experiments, there were no restrictions on the vehicle's movement and road conditions. The vehicle was driven freely as usual driving in the natural road conditions. The frame rate of video camera is 30 fps. There are 1,500 frames composing 50 sets of traffic signs in different locations. Each set contains 30 frames of image sequences. Each image is resized to 180 x 240 pixels.

To evaluate the tracking performance, the false-tracking rate defined as the ratio of the numbers of false tracked signs (signs are recognized, but they are tracked as the wrong objects or new objects) to the total numbers of signs in a set of image sequences (30 successive frames), is calculated for all 50 sets of image sequences. We compared the proposed method with two primitive methods: PM-1 and PM-2. PM-1 uses the two-layered blobs and employs the ring-partitioned matching as used by the proposed method, for the blob matching. However, the matching is performed for blobs in two successive frames. In PM-2, blob layer-1 and layer-2 are treated as a single blob and matched directly. The matching is performed for the blobs in the current frame and the ones in the temporary memory, the same to the one of the proposed method.

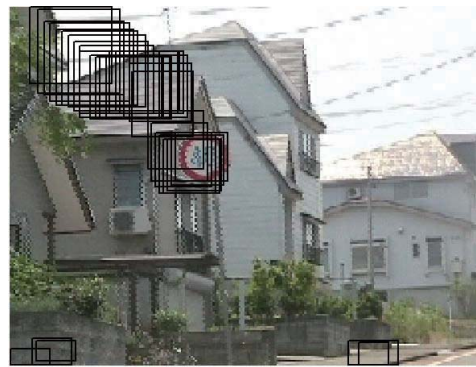
From the experiment, the false tracking rates of PM-1, PM-2, and the proposed approach are 7.2%, 5.5%, and 1.7%, respectively. This result showed the effectiveness of the temporary memory and the two-layered blob tracking proposed in the system.

To show the superiority of our approach from the Kalman-Filter tracking to overcome the problem of mis-detected signs during tracking, we carried out the experiment with the Kalman-Filter tracking as used in [3]. Fig. 3 (a) illustrates the trajectory of the sign tracked from a moving camera using our proposed method. Fig. 3 (b) illustrates the predicted y -coordinates of the center of the sign obtained by the Kalman-Filter tracking. In the figure, the predicted and the measured y -coordinates of the center of the signs are denoted with "dot" (\bullet) and "asterisk" ($*$), respectively. The bars drawn with circles (o) represent the predicted vertical sizes of the sign. The bars drawn with squares (\square) represent the measured vertical sizes of the signs.

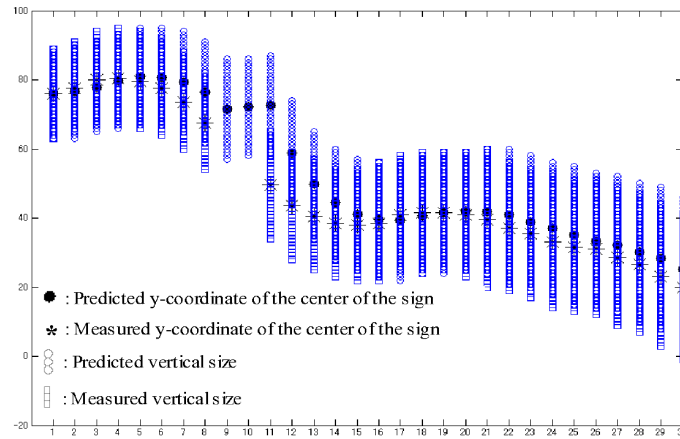
In Fig. 3 (b), sign is temporary mis-detected in frame-9 and frame-10. In the case of Kalman-Filter tracking, the predicted position and size are computed and used to predict the position and size of the sign in frame-11 in which sign is detected again. However, the predicted area in frame-11 is far away from the actual/measurement area. Thus, sign is not detected in frame-11. In contrast, our proposed tracking could detect sign properly as illustrated in Fig. 3 (a), because the search region is around the measured area (no prediction is made) and by introducing the temporary memory.

In order to examine the effects of tracking in reducing computation of the signs detection process, we compared the detection time of two cases; one is the case where traffic signs are detected from a single image, and the other is the case when utilizing the proposed tracking. In the former case, which is called WOT (a method without tracking), the detection technique [10] is applied to all the blobs obtained in the image. In the later case, called WIT (a method with tracking), the same detection technique is applied to the tracked blobs only. Both WOT and WIT use 50 images chosen arbitrary from 50 sets of image sequences, one image from each set.

The result of the detection times using MATLAB running on PC Intel Pentium-4, 2 GHz is 3.27second for WOT and 1.30second for WIT. From the result, the effect of tracking in reducing computation of the detection



(a)



(b)

Fig. 3. (a) Trajectory of traffic sign tracking; (b) Predicted and measured y-coordinate by the Kalman-Filter tracking.

process is clearly shown.

5 Conclusion

The system proposed a new approach of blob tracking to track traffic signs from image sequences. The tracking improved the performance of traffic sign detection by reducing computation time.

The proposed traffic sign tracking has advantages as follows:

- Even in the case of temporal misdetection, the sign could be tracked properly after it re-appears again;
- There is no need to make the model (accurate model) of the system,
- The camera could be installed on the vehicle easily, and there is no need to consider about the inertial navigation system of the vehicle,
- The vehicle could move freely or naturally, there is no need to restrict the vehicle movement, such as constant speed, straight movement, etc.

Since blobs are extracted by the color thresholding technique, selection of the color threshold values becomes a crucial task, due to the varying illumination and poor quality of image sequences. Our future research will address the improvement of the blob extraction technique.