International Journal of Innovative Computing, Information and Control
Volume **15**, Number **2**, April **2019**
 ISSN 1349-4198
pp. **653–665**

IMPROVING TRAFFIC SIGN DETECTION BY COMBINING MSER AND LUCAS KANADE TRACKING

Aryuanto Soetedjo and I Komang Somawirata

Department of Electrical Engineering
National Institute of Technology (ITN) Malang
Jalan Raya Karanglo KM 2, Malang 65143, East Java, Indonesia
aryuanto@gmail.com; kmgsomawirata@yahoo.com

Received May 2018; revised September 2018

Abstract. *This paper presents the combination of Maximally Stable Extremal Region (MSER) and Lukas Kanade Tracking (LKT) for detecting traffic sign. The proposed approach employs the MSER to find the red circular sign in an image. Once the traffic sign is detected, it is tracked by LKT to predict the position in the next frame. The detected sign and predicted position are treated as the traffic sign candidates. To reduce the false positive, the traffic sign candidates are validated using a template matching technique based on the Histogram of Oriented Gradient (HOG) feature. Instead of using MSER and LKT individually, our approach combines both methods to find the best candidate according to the matching score obtained by the validation stage. The experimental results show that our proposed method provides the efficient detection with the Recall of 0.9353, the Precision of 0.9747, the Accuracy of 0.9331, and the frame rate of 47.950 fps.*
Keywords: Traffic sign detection, MSER, Lukas Kanade tracking, HOG

1. Introduction.

An automatic traffic sign detection system based on the camera is one of popular and challenging topics in the intelligent transportation system. It detects the position of the traffic sign in an image for the further process such as sign classification or recognition. The traffic sign recognition is adopted for several applications [1,2]. a) It assists the driver to recognize the traffic signs properly when the driver is in the fatigue condition or when the complicated traffic signs appear on the road. By recognizing the traffic signs automatically, it helps the driver to avoid the missing signs. b) It is used by an autonomous vehicle for the navigation on the road. c) It could be used to collect the information of the physical condition of traffic signs for the maintenance purpose.

Traffic sign detection could be divided into several methods [3,4]: color-based methods, shape-based methods, the combination of color and shape methods, and Maximally Stable Extremal Region (MSER) method. In the color-based method, the color spaces are important aspects that should be considered. In the shape-based methods, the Hough transform, the ellipse/circle detection method, and Histogram of Oriented Gradient (HOG) features are usually employed. The MSER method originally proposed by [5] exploits the property of regions in which the shape remains the same when the threshold value is varied.

To improve the accuracy and the efficiency of sign detection, the tracking technique could be employed [6]. By tracking the sign, the position of the sign in the next frame could be predicted. The search area for detecting the sign is limited to the predicted position only. Thus, it reduces the computation time. Moreover, since the detected traffic sign is tracked from frame to the next frame, the false detection could be reduced.

DOI: 10.24507/ijicic.15.02.653

Due to the requirement of the real-time application, the computation time of the algorithm should be fast enough, while the accuracy should be high. Generally, the accuracy could be improved by employing the complex algorithms. However, the complex algorithms tend to consume the computation time.

In this paper, we propose a novel approach to improve the traffic sign detection accuracy and computation time by combining the traffic sign detection and tracking. In this work, we deal with the red circular signs. However, the approach could be extended to cope with the other types of traffic signs accordingly. In the detection stage, the MSER technique is adopted to detect the sign. A new color space [7] is adopted in the MSER process. After the detection stage, a simple efficient sign validation technique using HOG feature [4] is introduced to discard non-traffic sign objects which are detected wrongly by the detection method. This approach reduces the false detection effectively. In the tracking stage, the Lucas Kanade Tracking (LKT) is employed to track the sign. To improve the true detection rate, we propose a novel approach to combine the detection results obtained from the MSER and the LKT methods. The final detected traffic sign is selected according to the matching score of candidate obtained from the validation process.

The main contributions of our work are in the combination of detection and tracking techniques to avoid the misdetection, and the introduction of the sign validation process to discard the non-traffic sign objects. Further, the algorithm offers the fast processing time.

The rest of paper is organized as follows. Section 2 describes the literature review of the related work in the traffic sign detection and tracking. Our proposed detection and tracking is discussed in Section 3. Section 4 describes the experimental results. Finally, the conclusions are covered in Section 5.

2. Related Work.

2.1. Traffic sign detection. The Hue Saturation Intensity (HSI) color space was employed by [1] to detect the traffic sign. They proposed the color transformation of hue and saturation components by creating the look-up tables. Instead of using a fixed threshold, the look-up table is adopted to avoid the segmentation error. Both components work complementary, i.e., when the hue component fails to detect the object, the saturation component will correct it, and vice versa.

The red and blue color thresholding based on RGB color space was proposed by [5] to detect the red and blue colors of the traffic sign. A pixel is assigned as a red object if the following condition is satisfied:

$$r \geq g \text{ and } r \geq b \text{ and } g/|r - g| \leq 2.5 \tag{1}$$

where r, g, b are the normalized red, green, and blue components of a pixel, respectively. A pixel is assigned as a blue object if the following condition is satisfied:

$$g/b \leq 0.65 \tag{2}$$

To detect the red traffic sign, the difference between the normalized red component and the normalized green component was employed [7]. The method assigns a pixel as a red object if the following condition is satisfied:

$$g - r \leq Thr \tag{3}$$

where Thr is a threshold which is determined automatically by observing the g-r histogram. The g-r histogram is a histogram which is developed by counting the number of pixels based on the value of $g - r$ (subtracting the normalized red from the normalized green). Once the histogram is calculated, the peak-valley analysis is used to find the threshold *Thr*.

The traffic sign detection based on the MSER was proposed by [8]. The method detects the traffic sign according to the background color. For the traffic sign with the white background, it is first converted to the grayscale image. Then the MSER method is applied to the grayscale image. To reduce the false detected objects, the selection based on the geometry properties such as the width, height, aspect ratio, region perimeter and area, and bounding-box perimeter and area is employed. For the traffic sign with the red or blue background, the image is converted to the normalized red/blue image using the following formula:

$$Norm_{RB} = \max\left(\frac{R}{R+G+B}, \frac{B}{R+G+B}\right) \tag{4}$$

where R, G, B are the red, green, blue components in the RGB color space. This color conversion was also adopted by [9]. The MSER method is insensitive to the illumination changes. However, the computation time is high. To reduce the computation time, the thresholds are limited to a certain range only [8].

2.2. **Traffic sign tracking.**

2.2.1. *Kalman filter tracking.* The common method to track the traffic sign is using the Kalman filter. The Kalman filter requires the state model and the measurement model for prediction and correction processes. The state model is expressed in the following equation:

$$x_t = F_t x_{t-1} + w_t \tag{5}$$

where x_t is the state vector at time t, F_t is the state transition matrix at time t, and w_t is the model noise. The measurement model is expressed as

$$z_t = H_t x_t + v_t \tag{6}$$

where z_t is the measurement vector at time t, H_t is the state transformation matrix at time t, and v_t is the measurement noise. The Kalman filter was employed to predict the position and size of the traffic sign [10]. The method assumes that the vehicle's speed and the size of the traffic sign are known. The state vector is the position of the traffic sign in horizontal and vertical directions, and the radius of the traffic sign in the previous frame. The measurement vector is the position of the sign in horizontal and vertical directions, and the radius of the sign in the current frame. The similar traffic sign tracking method was proposed by [11].

The Kalman filter proposed by [12] is similar to the previous one by extending the variables of the state vector. In addition to the position of the sign in horizontal and vertical directions, and the radius of the sign, their derivatives are included in the state vector. The works in [13,14] used the position of the traffic sign in the horizontal and vertical directions, and the width and the height of the traffic sign in the state vector and the measurement vector. To improve the stability of the model, the state vector is split into two state vectors, i.e., the location state vector and the size state vector [14]. The traffic sign tracking proposed by [15] used two state vectors, i.e., the vector represents the position and velocity of the traffic sign, and the vector represents the width and the height of the traffic sign.

2.2.2. *Lukas Kanade Tracking (LKT).* Traffic sign tracking using LKT was employed by [16-18]. The LKT method is described as follows [19]. Let $I(x, y, t)$ be the intensity of pixel (x, y) at time t. Under the assumption that the intensity does not change in the consecutive frames and the pixels in a small region have a similar motion, then it is obtained that

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + 1) \tag{7}$$

where δx, δy are displacements along horizontal and vertical directions respectively. Using the Taylor expansion the following constraint is obtained as

$$I_x u + I_y v + I_t = 0 \tag{8}$$

where

$$I_x = \frac{\delta I}{\delta x}; \quad I_y = \frac{\delta I}{\delta y}; \quad I_t = \frac{\delta I}{\delta t}; \quad u = \frac{dx}{dt}; \quad u = \frac{dy}{dt} \tag{9}$$

The LKT method solves the above equation using the least square fit method as given in the following

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_i I_{xi}^2 & \sum_i I_{xi} I_{yi} \\ \sum_i I_{xi} I_{yi} & \sum_i I_{yi}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i I_{xi} I_{ti} \\ -\sum_i I_{yi} I_{ti} \end{bmatrix} \tag{10}$$

where index i denotes the number of pixels in a small window.

2.2.3. *Other tracking methods.* The work in [20] predicted the position of the traffic sign in the 3rd frame using the information obtained from the 2nd and 1st frames. The method assumes that the speed of the vehicle is constant during the detection process. The horizontal position of the traffic sign in the 3rd frame (x_3) is calculated using the following formula

$$x_3 = \frac{x_1 x_2}{\left(1 + \frac{t_2}{t_1}\right) x_2 - \frac{t_2}{t_1} x_1} \tag{11}$$

where x_1, x_2 are the horizontal positions of the traffic sign in the 1st and 2nd frames respective, and t_1 and t_2 are the times to move from the 1st to 2nd frame and the 2nd to 3rd frames respectively.

The work in [21] proposed a traffic sign tracking based on a stereo vision. The position of the traffic sign in the next frame is tracked based on the vehicle's motion which is calculated by the stereo vision technique. To compensate for the prediction error, the template matching is applied in the regions surrounding the predicted position. Blob tracking was employed by [22] to improve the performance of traffic sign detection. The method matches the blobs in the current frame to the ones in the previous frames. Since the latest tracked blobs are stored in a memory, it could recover the temporarily disappeared blobs. The ring partitioned matching is adopted to match the different sizes of the blobs in the successive frames.

3. **Proposed Method.**

3.1. **Overview of the proposed method.** Block diagram of the proposed method is depicted in Figure 1. In the figure, the left part shows the process in frame-k, while the right part shows the process in frame-$(k + 1)$. It is assumed that in frame-k the traffic sign is detected for the first time. Then it is tracked in frame-$(k + 1)$ and so on. In the detection stage, the MSER method is employed. Since the MSER method varies the threshold to find the stable region, it provides an effective object segmentation. However, it is time-consuming. To reduce the computation time, a method to limit the search area is introduced as shown in Figure 1.

It is common that the detected objects obtained by the detection stage may contain the non-traffic sign objects. Therefore, a validation stage is required to discard them. The validation technique should be the fast and effective algorithm. In this work, we adopt

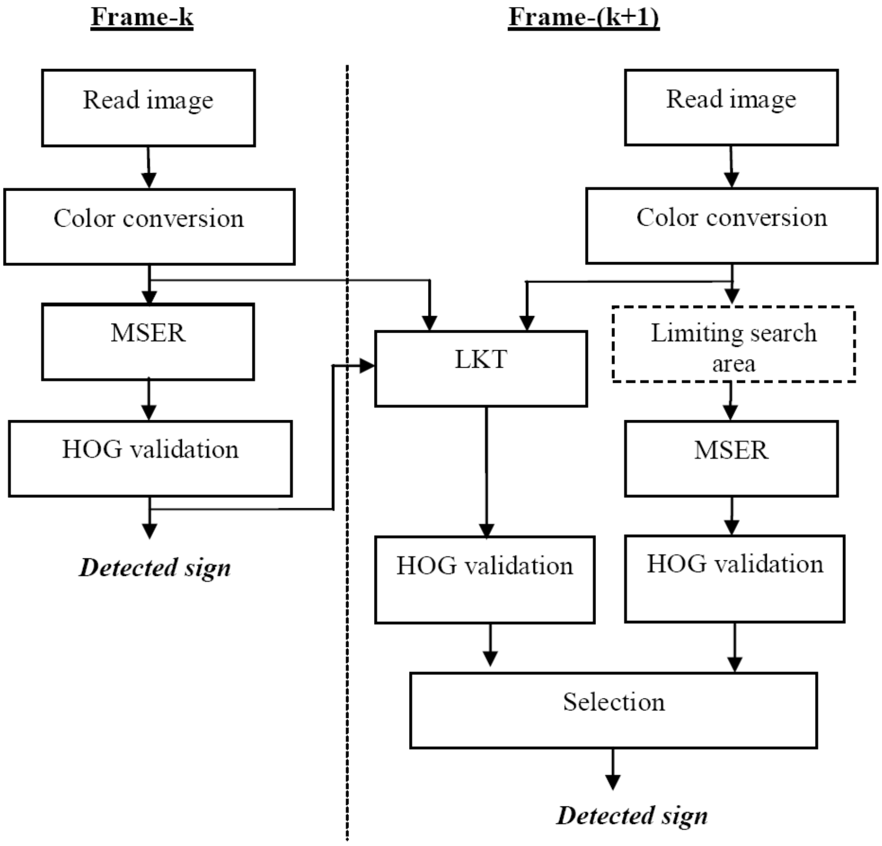
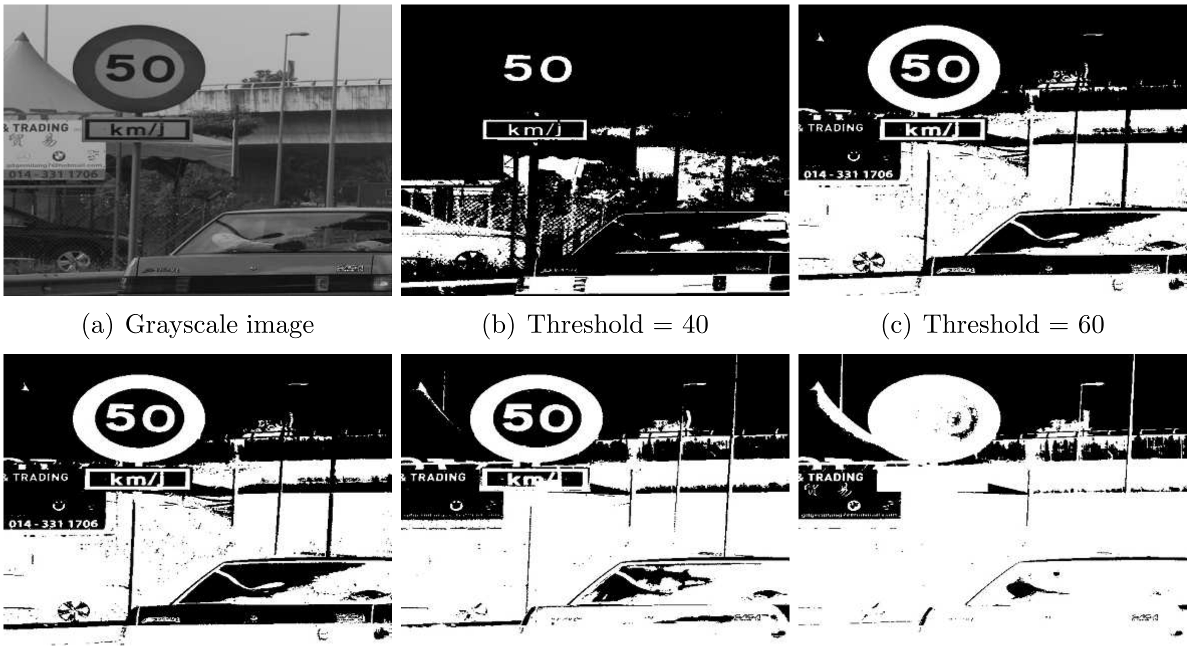


FIGURE 1. Block diagram of the proposed method

the validation technique using the template matching based on HOG feature [4], which is described in Section 3.3.

Once the traffic sign is detected in a frame, it is tracked using the LKT. As discussed previously, the LKT provides an effective tracking without modeling the vehicle's movements, which is impractical in the real implementation. The method first finds the points to be tracked in frame-k. Then the corresponding points in frame-$(k+1)$ are searched using the formulas given in (7)-(10). The tracked points are the corners of the bounding box of the detected traffic sign. Generally, the tracked bounding box is used to limit the search area of detection. In our approach, in addition to limiting the search area, we also treat the tracked bounding box as the candidate detected traffic sign together with the one obtained by the detection stage. This approach increases the true detection which is described clearly in Section 3.4. Finally, the detected sign is selected from the candidates according to their validation scores.

3.2. **Traffic sign detection using MSER.** The MSER finds the stable region in an image by varying the threshold in the segmentation process [5]. The idea of MSER method is illustrated as follows. If we apply the thresholding technique on a grayscale image at the different threshold levels, from the low value to the high value, the results of thresholded images are depicted in Figure 2. The connected regions (represented by white color) start from a small size at the low threshold value to a large size, i.e., occupy a whole image, at the high low threshold value. The extremal regions are the set of connected regions in the sequence of thresholded images. In the MSER method, the maximally stable extrema regions are found when the shapes of extremal regions are nearly the same over the variation of thresholds. From Figure 2 we can see that the regions of the number



(a) Grayscale image (b) Threshold = 40 (c) Threshold = 60

(d) Threshold = 80 (e) Threshold = 100 (f) Threshold = 120

FIGURE 2. Image thresholding under the different threshold levels

of 50, circular ring of the traffic sign, the wall of the bridge remains the same over several threshold values. It leads to the motivation that the MSER method can be used to detect the traffic sign effectively.

Since the MSER method works on the grayscale image, the color image should be converted into the grayscale image. Since this work deals with the red circular sign and inspired by our previous work [7], instead of using the color conversion as proposed by [8,9], we employ the g-r color conversion which provides the effective red color segmentation.

3.3. **Traffic sign validation.** The traffic sign validation is employed to effectively discard the non-traffic sign objects obtained from the detection stage [4]. Since it needs an additional time, the validation method should be fast. Our method employs the template matching technique using the HOG feature of the image as described in the following.

The HOG feature is first introduced by [23]. It represents the histogram of the orientation of the gradient in an image. The histogram is computed as follows:

- Resize an image into a specific size (i.e., 40×40 pixels in this work)
- Divide the image into 8×8 cells, where each cell is 5×5 pixels (see Figure 3)
- Define the overlapping block which covers 2×2 cells (see Figure 3)
- Compute the magnitude and orientation of the gradient of each pixel in the block
- Quantize the orientation into the bins (i.e., 9 bins in this work)
- Accumulate the weighted votes for the magnitudes into the bins of the pixels over the cell
- Normalize the contrast of the block
- Concatenate the histogram of all overlapping blocks

Let HC be the histogram of each cell, and then the histogram on a block is $HB = (HC_1, HC_2, HC_3, HC_4)$, where HC_1 to HC_4 is the histogram of cell-1 to cell-4 of the block. The final histogram of the image is $HOG = (HB_{11}, HB_{12}, \ldots, HB_{ij}, \ldots, HB_{77})$, where HB_{ij} is the histogram on the block of the i-th position in the horizontal direction and j-th in the vertical directions. It is noted here that since there are 8×8 cells in the

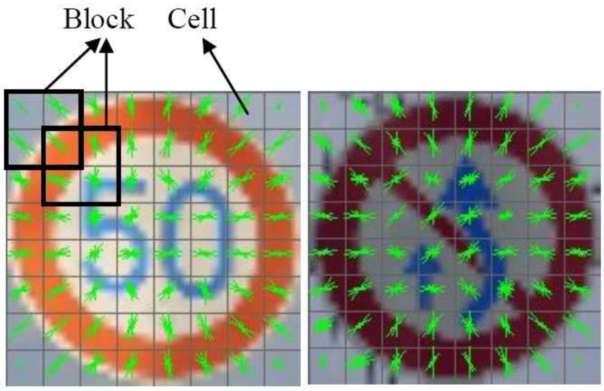


Figure 3. Visualization of HOG feature

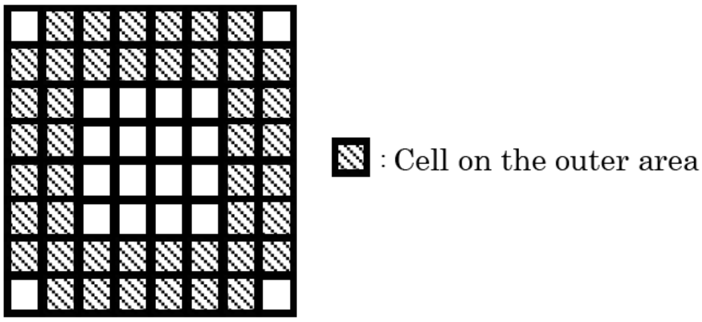


Figure 4. Outer area for HOG calculation

image, the number of overlapping blocks is $7 \times 7 = 49$. Since the number of bins is 9, then the final histogram (HOG) has $49 \times 4 \times 9 = 1764$ dimensional vector (i.e., the number of blocks $\times$ the number of cells on each block $\times$ the number of bins).

The visualization of the HOG of traffic sign image is illustrated in Figure 3, where the green lines drawn in each cell (small window) denotes the orientation of the image in the cell. By observing Figure 3, it is obtained that the HOG features of two different traffic signs in the outer area (red circular ring) show a similar orientation. It could be understood from the fact that the orientation of the circular ring is the same along its perimeter. Since the validation stage is used to verify that the detected sign is a red circular sign (not for sign classification), we may use the property of HOG feature in the outer ring for the red circular traffic sign validation. The simple way is by applying the template matching based on the HOG feature. Instead of applying the HOG of a whole image, our method computes the HOG feature on the outer area only as shown in Figure 4. Let $THOG$ and $RHOG$ be the HOGs of the test image and reference image respectively. The distance between the test image and the reference image ($Dist\_TR$) is given as

$$Dist\_TR = \sum_{i,j \in outer} (THOG_{ij} - RHOG_{ij})^2 \tag{12}$$

where i, j are the horizontal and vertical positions of the block as defined previously, and $outer$ is the outer area of the image as shown with the shaded boxes in Figure 4. The test image is validated as the circular traffic sign if the value of $Dist\_TR$ is lower than a threshold.

3.4. **Combining detection and tracking.** Our novel approach to combining the detection and tracking is based on the observation that both the tracking method and the detection method have the complementary benefits as discussed in the following.

The tracking technique is commonly employed to improve the detection performance. Once the sign is detected, its position in the next frame is predicted by the tracker. The predicted position is then used to limit the search area of sign detection. Since the detection stage is applied in the limited area, it has two benefits: a) reduce the detection time; b) reduce the false detection. However, in some conditions, the quality of images in a video may degrade due to the noises or the camera's vibration. In this situation, the detection stage may fail to detect the traffic sign.

Another approach to utilizing the tracking technique is by treating the tracked position as the detected sign. In this approach, the detection stage is employed only in the first attempt to detect the sign. Once the sign is detected, the tracking technique is applied to track (detect) the sign in the consecutive frames. Since the execution time of the tracking technique is usually faster than the detection technique, the approach reduces the computation time effectively. The limitation of the approach is when the new traffic sign appears in the image. The sign will not be detected by the tracker.

Our proposed approach exploits the benefits of both the detection and tracking techniques by applying both techniques in every frame while treating the predicted position of the tracker as the candidate sign. In this approach, the sign candidate may come from three different situations: a) the sign candidate is detected by the detection technique only; b) the sign candidate is detected by the tracking technique only; c) the sign candidates are detected by the detection and tracking techniques. In the last case, a candidate selection method should be developed to select the best candidate.

Our approach to selecting the best traffic sign candidate takes a benefit of the proposed HOG validation process, in which each candidate obtained by the validation stage has a score that denotes the matching degree, which is expressed as $Dist\_TR$ in (12). The candidate with the best score, i.e., the lowest $Dist\_TR$ is selected as the final detected sign. In the case of multiple traffic signs in an image, the candidates are grouped according to their positions in the image.

4. **Experimental Results.** To verify our proposed method, we test the algorithm using the real video images. The video images are taken from a video camera installed inside a car. The video camera is the Canon EOS 60D with the video resolution of 1920×1080 pixels. The traffic sign images are taken on several streets in Malang city, Indonesia during daytime (morning until afternoon). There are 94 video datasets that contain the red circular traffic sign images on the streets. Some images grabbed from the tested videos are illustrated in Figure 5. In the experiments, the images are resized into 800×480 pixels. The algorithm is implemented on a personal computer Intel Core i7 3.4 GHz. The C++ software language and OpenCV library are employed to test the algorithm.

In the experiments we compute the Recall, Precision and Accuracy which are expressed as follows:

$$Recall = \frac{TP}{TP + FN} \tag{13}$$

$$Precision = \frac{TP}{TP + FP} \tag{14}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

where TP is the true positive, FN is the false negative, FP is the false positive, and TN is the true negative.

Five scenarios are examined: a) The MSER using color conversion as proposed by the existing methods [8,9], which is called as **MSER-Ext**; b) The MSER using color conversion as proposed in our previous methods [7], which is called as **MSER-Prop**;

FIGURE 5. Example of static images grabbed from the tested videos

TABLE 1. Experimental results of Recall, Precision, and Accuracy

Method	**Recall**	**Precision**	**Accuracy**
MSER-Ext	0.7722	0.8120	0.7192
MSER-Prop	0.8191	0.9167	0.8144
LKT	0.8678	0.9576	0.8710
Combination	0.9361	0.9136	0.8867
Combination-Limit	0.9353	0.9747	0.9331

c) The LKT method, which is called as **LKT**; d) Combination of **LKT** and **MSER-Prop**, which is called as **Combination**; e) Combination of **LKT** and **MSER-Prop** with limiting the search area, which is called as **Combination-Limit**. In this case, the limited search area is defined as the region in which the area is 1.5 times of the previous detected bounded area.

The experimental results are given in Table 1, where the values of Recall, Precision, and Accuracy are the average values from 94 video datasets. From the table, it is obtained that the proposed MSER using color conversion developed in our previous work [7] is superior to the ones of existing methods [8,9] as indicated by the higher Recall, Precision, and Accuracy. The Recall and Accuracy of combination methods are higher than the individual methods. The results can be understood from the fact that by combining two methods, the true positive (TP) will increase due to the complementary effect, i.e., when the traffic sign is not detected by the LKT, it may be detected by MSER, and vice versa. According to (13), since $TP+FN$ is fixed for all methods, the Recall will increase when TP is increased. The similar condition applies for the Accuracy as given in (15), i.e., since $TP+FN+TN+FP$ is fixed and the combination method does not affect the TN, the Accuracy will increase when TP is increased. The result shows that the combination method increases the Recall significantly, i.e., from 0.8678 (**LKT**) and 0.8191 (**MSER-Prop**) to 0.9361 (**Combination**). The Recall of the combination method with limiting the search area is slightly lower than the one without limiting the search area. It is caused by some traffic signs that are not detected due to the limited search area.

From Table 1, it is obtained that the Precision of **Combination** (combination method without limiting the search area) is lower than **LKT** and **MSER-Prop**. It is clear that by combining the LKT and MSER, both the TP and FP will increase. Thus according to (14), the Precision tends to decrease. Fortunately, by limiting the search area, the FP will decrease. Therefore, the Precision will increase as shown by the highest Precision of **Combination-Limit**.

To examine the computation time of the proposed method, the color conversion time, the execution time of detection methods, and the total frame rates are measured as given in Table 2. From the table, it is obtained that the color conversion time of our work is faster than the existing methods [8,9]. Moreover, the detection time of our proposed combination method is decreased significantly by limiting the search area that yields a high frame rate of 47.950 fps (**Combination-Limit**).

Table 2. Results of execution time

Method	Color conversion time (ms)	Detection time (ms)	Frame rate (fps)
MSER-Ext	20.451	19.018	25.336
Combination	10.137	23.989	29.303
Combination-Limit	10.137	10.718	47.950

The time complexity of the proposed algorithm is discussed in the following. The algorithm consists of the color conversion step and the detection step. The color conversion time is affected by the number of pixels in the image. Thus, the complexity of color conversion is $O(N)$ where N is the number of pixels in the image. In the detection step, the time complexity of MSER method is $O(N\log(\log(N)))$ [5]. While the time complexity of the LKT is $O(n^2N + n^3)$ [24], where n is the number of warp parameters ($n = 2$ in this work). Therefore, the time complexity of our proposed method is $O(N + N\log(\log(N)) + n^2N + n^3)$. It suggests that the number of pixels in the image (N) is the dominant parameter that affects the time complexity, more specifically in the MSER detection method. It is clearly shown that by limiting the search area (**Combination-Limit**), the number of N is reduced, and thus it reduces the time complexity as given in Table 2.

From the above results, we may conclude the superiority of our approach as follows. The strategy to combine the MSER and LKT will increase the Recall and Accuracy by increasing the TP. Meanwhile, the combination method increases the FP that decreases the Precision. To overcome this drawback, we introduce the strategy to limit the search area. By limiting the search area, the FP is decreased that increases the Precision. Further, since the search area is limited, the detection time will be decreased significantly.

The samples of detection results are illustrated in Figure 6, where the blue square denotes the result of **Combination-Limit**, the green ellipse denotes the result of **MSER-Prop**, and the yellow rhombus denotes the result of **LKT**. As shown in the figure, in the 39^{th} frame, **MSER-Prop** fails to detect the traffic sign. However, **LKT** and

(a) Video dataset-10, frame = 39

(b) Video dataset-10, frame = 40

(c) Video dataset-10, frame = 41

FIGURE 6. Detection results in video dataset-10. Blue square (**Combination-Limit**), green ellipse (**MSER-Prop**), yellow rhombus (**LKT**): (a) frame = 39; (b) frame = 40; (c) frame = 41.

Combination-Limit detect the traffic sign properly. In the 40$^{\text{th}}$ frame, **LKT** fails to detect the traffic sign. However, **MSER-Prop** and **Combination-Limit** detect the traffic sign properly. In the 41$^{\text{st}}$ frame, all three methods detect the traffic sign properly.

5. **Conclusions.** A method to improve the traffic sign detection is presented. The key points of the proposed method are threefold. Firstly, it introduces the HOG validation method to reduce the false position detection. Secondly, it combines the MSER detection

and the LKT to increase the true positive detection. Thirdly, it employs a simple method to limit the search area to speed up the computation time and reduce the false positive detection. The effectiveness of the proposed method is proved by the experiments that achieve the high values of Recall, Precision, Accuracy and the fast computation time.

In future, some improvements in the detection stage and validation stage will be conducted by adopting the new and robust techniques. The method will be extended to handle the other traffic sign images. Furthermore, the implementation on an embedded system will be carried out.

Acknowledgment. This work is supported by the Research Grant, Competence-Based Research scheme from Directorate General of Higher Education, Ministry of Research and Technology and Higher Education, Republic of Indonesia (SP DIPA. 042.06.1.401516/ 2018).

REFERENCES

[1] 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, 2003.
[2] A. Gudigar, S. Chokkadi and R. U, A review on automatic detection and recognition of traffic sign, *Multimedia Tools and Applications*, vol.75, no.1, pp.333-364, 2016.
[3] S. K. Berkaya, H. Gunduz, O. Ozsen, C. Akinlar and S. Gunal, On circular traffic sign detection and recognition, *Expert Systems with Applications*, vol.48, pp.67-75, 2016.
[4] A. Soetedjo and I K. Somawirata, An efficient algorithm for implementing traffic sign detection on low cost embedded system, *International Journal of Innovative Computing, Information and Control*, vol.14, no.1, pp.1-14, 2018.
[5] J. Matas, O. Chum, M. Urban and T. Pajdla, Robust wide-baseline stereo from maximally stable extremal regions, *Image and Vision Computing*, vol.22, no.10, pp.761-767, 2004.
[6] A. S. Lafuente, B. S. Maldonado, J. P. Gil, R. J. Acevedo and S. R. J. Lopez, A tracking system for automated inventory of road signs, *Proc. of IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey, pp.166-171, 2007.
[7] A. Soetedjo and K. Yamada, A new approach on red color thresholding for traffic sign recognition system, *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, vol.19, no.4, pp.458-465, 2007.
[8] J. Greenhalgh and M. Mirmehdi, Real-time detection and recognition of road traffic signs, *IEEE Trans. Intelligent Transportation Systems*, vol.13, no.4, pp.1498-1506, 2012.
[9] L. Kurnianggoro, Wahyono and K.-H. Jo, Improving traffic sign recognition using low dimensional features, *Proc. of Asian Conference on Intelligent Information and Database Systems*, Kanazawa, Japan, pp.237-244, 2017.
[10] C. Y. Fang, S. W. Chen and C. S. Fuh, Road-sign detection and tracking, *IEEE Trans. Vehicular Technology*, vol.52, no.5, pp.1329-1341, 2003.
[11] G. M. A. Garcia, M. A. Sotelo and G. E. Martin, Fast traffic sign detection and recognition under changing lighting conditions, *Proc. of IEEE Intelligent Transportation Systems Conference*, Toronto, Canada, pp.811-816, 2006.
[12] G. M. A. García, M. Ocana, D. F. Llorca, E. Arroyo, J. Pozuelo and M. Gavilan, Complete vision-based traffic sign recognition supported by an I2V communication system, *Sensors*, vol.12, no.2, pp.1148-1169, 2012.
[13] A. H. Mazinan and M. Sarikhani, Providing an efficient intelligent transportation system through detection, tracking and recognition of the region of interest in traffic signs by using non-linear SVM classifier in line with histogram oriented gradient and Kalman filter approach, *Sadhana*, vol.39, no.1, pp.27-37, 2014.
[14] K. Wang, Z. Hou and W. Gong, Automated detection, tracking, and recognition of roadway signs, *Proc. of the 1st Annual Inter-University Symposium on Infrastructure Management*, Waterloo, Canada, pp.1-18, 2005.
[15] A. Ruta, Y. Li and X. Liu, Detection, tracking and recognition of traffic signs from video input, *Proc. of the 11th International IEEE Conference on Intelligent Transportation Systems*, Beijing, China, pp.55-60, 2008.

[16] D. Juric and S. Loncaric, Warning and prohibitory traffic sign detection based on edge orientation gradients, *Proc. of the Croatian Computer Vision Workshop*, Zagreb, Croatia, pp.27-32, 2014.

[17] E. Krsak and S. Toth, Traffic sign recognition and localization for databases, *Acta Electrotechnica et Informatica*, vol.11, no.4, pp.31-35, 2011.

[18] N. B. Romdhane, H. Mliki and M. Hammami, An improved traffic signs recognition and tracking method for driver assistance system, *Proc. of IEEE/ACIS the 15th International Conference on Computer and Information Science*, Okayama, Japan, pp.1-6, 2016.

[19] B. D. Lucas and Kanade, An iterative image registration technique with an application to stereo vision, *Proc. of DARPA Image Understanding Workshop*, Vancouver, Canada, pp.121-130, 1981.

[20] J. Miura, T. Kanda, S. Nakatani and Y. Shirai, An active vision system for on-line traffic sign recognition, *IEICE Trans. Information and Systems*, vol.85, no.11, pp.1784-1792, 2002.

[21] C.-W. Choi, S.-I. Choi and S.-Y. Park, Efficient detection and tracking of road signs based on vehicle motion and stereo vision, *Proc. of International Conference on Advanced Concepts for Intelligent Vision Systems*, Poznań, Poland, pp.608-619, 2013.

[22] A. Soetedjo and K. Yamada, Improving the performance of traffic sign detection using blob tracking, *IEICE Electronic Express*, vol.4, no.21, pp.684-689, 2007.

[23] N. Dalal and B. Triggs, Histograms of oriented gradients for human detection, *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, USA, pp.886-893, 2005.

[24] S. Baker and I. Matthews, Lucas-Kanade 20 years on: A unifying framework, *International Journal of Computer Vision*, vol.56, no.3, pp.221-255, 2004.