

# Warning and Prohibitory Traffic Sign Detection based on Edge Orientation Gradients

Darko Jurić and Sven Lončarić

University of Zagreb, Faculty of Electrical Engineering and Computing,  
Department of Electronic Systems and Information Processing

Unska 3, 10000 Zagreb, Croatia

Email: {darko.juric, sven.loncaric}@fer.hr

**Abstract**—A method for traffic sign shape detection is presented, which does not require extensive training and is based on fast template matching suitable for real-time applications. The proposed method is able to detect prohibitory and warning signs. Post processing phase enables detection verification thus drastically increasing detection accuracy. All method parameters, if required, can be tweaked without the need for time-consuming retraining procedures. The method shows robustness to various conditions that often impose problems for color based systems. The proposed method can be used standalone or with already developed methods as a verification layer.

**Keywords**—computer vision, traffic sign detection, template matching

## I. INTRODUCTION

Traffic sign detection and recognition is an important problem in advanced driver assistance systems (ADAS). ADAS utilize sophisticated algorithms, which not only address traffic sign detection and recognition but other applications including lane detection, driver distraction detection, and parking assistance. Poor visibility and/or driver fatigue can be one of the main reasons that can lead to traffic accidents. While traffic sign detection systems are commercially available, they still present an open research topic. In recent years, speed limit detection systems have been included in top-of-the-line vehicle models from various manufactures, but these systems are still limited to a subset of all traffic signs. Therefore, there is a need for traffic sign detection and recognition methodology, which is capable of detecting a wider selection of traffic signs. Nowadays, embedded devices not only contain basic FPGA units, which can be designed to support vectorized array addition, but CPU support as well thus enabling much broader range of usage. The important feature is ability to tweak detection parameters without additional training. Many methods that are considered state-of-the-art do not have such possibility like [1] and [2]. The paper is organized as follows. A short overview of related work is presented in Sec. II. Sec. III describes the proposed method. Sec. IV gives used parameters and results, which are obtained on videos that contain challenging lightening conditions and extensive clutter. Conclusion is given in Sec. V.

## II. RELATED WORK

### A. Color-based techniques

Some parts of the method presented in [3] are based on simple thresholding to distinguish a particular color (red) from

others. However, various factors such as color change due to illumination factor (e.g. time of a day) or non-uniform illumination that produces shadows need more sophisticated methods like [4] in which a light source color is estimated. In [5] features are extracted after image segmentation process which is done by using Gabor filters, K-means clustering and smoothing filters. In [6] CIECAM97 color model is used where only hue and chroma channels are used. Various daylight viewing cases are considered. A general drawback of color based techniques is that they require special measures to achieve luminance invariance.

### B. Shape-based techniques

The most common technique for shape-based approaches is Hough transform. A method that uses a modified Hough transform to detect circular, triangle and square shaped signs is presented in [7]. Distance transform (DT) based techniques are another popular approach for traffic sign detection. In [8], a modified DT technique is used. First, edge orientations are extracted from the image, then a DT image is built for every edge orientation. A score match is obtained in a way that a template is rotated and scaled and matched against every DT image. The maximum score is taken as the score match. Hough transform is more constrained regarding shape diversities – it can approximate only simple shapes and traditional template searching can be time consuming.

### C. Machine learning based techniques

The most common used technique is based on work of Viola and Jones presented in [9]. Although it was originally developed for face-detection it can be successfully used for detection of other objects, including traffic signs. In [2], a traffic sign detection and recognition method is presented, which uses the mentioned technique. Although Viola-Jones approach is widely used it suffers from the following problems. A large amount of training samples is needed to train the detector. For some problems this imposes a significant obstacle. By introducing more object classes several detectors must be used in order to classify them. This can impose performance penalty.

## III. PROPOSED METHOD

In this paper, we present a novel method for traffic sign detection, which is based on fast edge orientation template matching [10]. The proposed method is robust to clutter and is computationally efficient and suitable for real-time

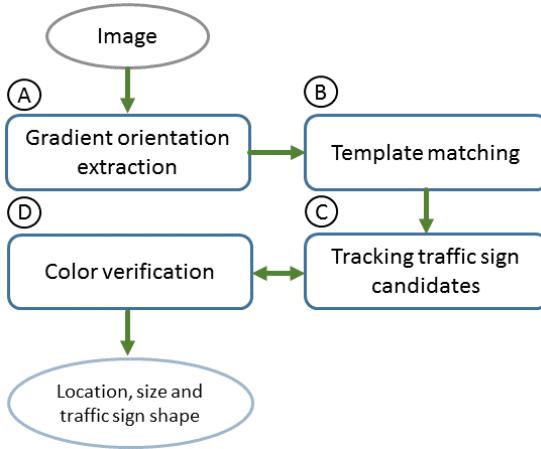


Fig. 1: A method composed of template matching as the core part and post processing stages which include shape verification as well as continuous tracking and detection consistency verification.

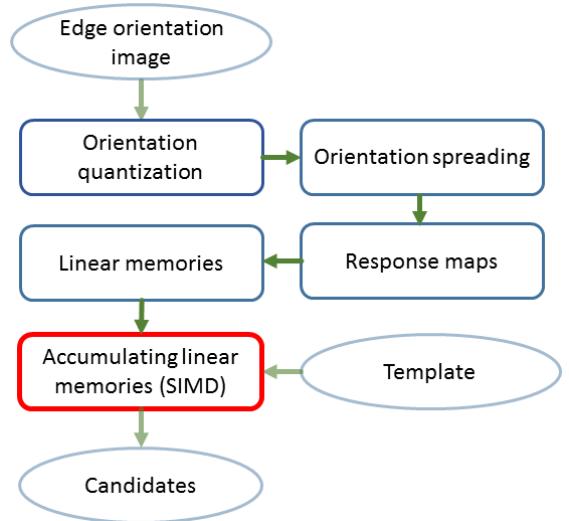


Fig. 2: Template matching pipeline. The red color highlights the modified part.

applications. The outline of the proposed method is shown in Fig. 1. The template matching algorithm [10] is modified in this paper in order to improve its accuracy by using a different gradient extraction method as described in Sec. III-A and by extending the maximum number of features per template, described in Sec. III-B1. However, the modified template matching procedure is still not sufficient for successful traffic sign detection due to higher false positive rate which is a result of an extensive clutter (e.g. forest along the road).

The post processing phase consists of two steps. The first step involves continuous traffic sign detection and tracking, described in Sec. III-C which reduces false positives considerably. The second phase eliminates structures that have similar shapes but do not have similar border color - validates traffic sign border-color (redness) described in Sec. III-D.

#### A. Gradient orientation extraction

The crucial preprocessing step is edge extraction, where an edge orientation image is an input image for the template matching method [10]. One of the main reasons that image orientations are chosen is their robustness to extensive light changes and image noise. Gradient magnitude performs poorly in scenes with background clutter due to many false positives. In [10], a multi-channel Sobel operator has been used where a pixel orientation is taken from the channel that has the largest magnitude. In addition, orientations are filtered in a way that only dominant orientations in a  $3 \times 3$  neighborhood are retained. Instead of the multichannel Sobel operator and orientation filtering a standard grayscale Canny implementation is used in the proposed method. By using this operator better results are obtained in terms of a reduced number of false positives in highly cluttered images (e.g. forest as background).

#### B. Template matching

After the gradient orientation extraction presented in Sec. III-A is performed, the modified template matching algorithm is used in order to detect circular and triangular shapes -

traffic sign candidates. The outline of the template matching procedure is shown in Fig. 2. Only relevant parts needed to understand the modifications are going to be briefly discussed. For full explanation of the procedure see [10].

The template matching procedure uses quantized edge orientations as features where each quantized orientation is encoded in 8-bit vector as a single bit flag. The template is matched against every location within the image using the operation equivalent to sliding a template over the input image. In contrast to the standard template matching procedure the input image is preprocessed so that the matching procedure can be executed very fast by adding long arrays - linear memories [10]. Each linear memory cell contains similarity in range  $[0..n]$ , where  $n$  is maximum user-specified similarity between two quantized orientations. These values are limited by the number of bits in 8-bit value. The linear memory addition can be done very efficiently by using SIMD processor architecture. The addition process is done as follows.

Each template feature, see Fig. 5, which consists of a location and its corresponding quantized orientation is used to select linear memory (8-bit vector). The selected memories are added to the 8-bit result vector initialized by zeroes. The result vector contains the raw template matching scores in range  $[0..255]$  for each location. The values are rescaled to the range  $[0..100]$ . The result vector is thresholded in order to extract only strong matches. For details about the linear memory content and creation see [10].

*1) Extending number of features:* The 8-bit result vector dictates the maximum number of features per template because each linear cell contain value which maximum is  $n$  and the number of linear memories being added corresponds to the number of template features. See Fig. 3 for details. Therefore the maximum number of features per template, in the original paper, is limited to  $\lfloor 255/n \rfloor$ . This number of features has not shown to be enough to robustly detect triangular and circular shapes in strong clutter such as forest. To overcome this limitation the linear maps are first added

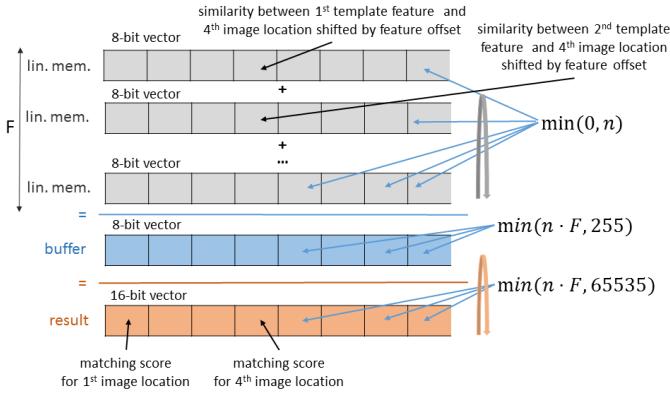


Fig. 3: The proposed procedure for linear memory addition - template matching. Linear memories are added to the buffer which is periodically added to the final result vector.  $n$  represents the maximum orientation similarity, and  $F$  is the number of features of a template.

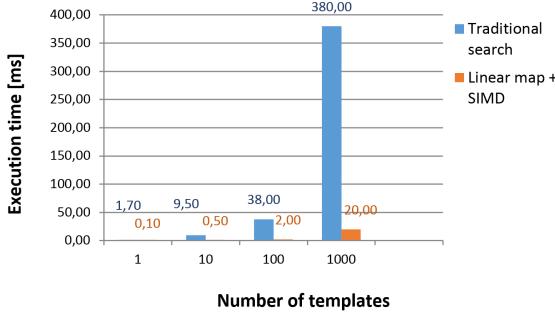


Fig. 4: Traditional and speeded-up template matching performance comparison. The number of features per template is 100. All tests are conducted on Intel Core i5-2500K CPU using single core.

to a temporary buffer - 8 bit vector, and then before the buffer overflows, the buffer is added to the final result array (16-bit), which contains template matching scores. See Fig. 3 for details. This procedure is repeated for each template feature. The buffer is cleared after its values are accumulated to result array. The result is the increased maximum number of features  $\lfloor 65535/n \rfloor$ . Fig. 4 shows the significant speed-up over traditional template matching approach.

**2) Building traffic sign templates:** Feature extraction can be performed directly on color images, but better results are obtained if a binarized image is used, which is first smoothed. The number of extracted features is fixed (it does not depend on an object size). At far distances Canny edge detector usually cannot detect outer and inner circle or triangle, which are the integral parts of the warning and prohibitory signs. Therefore, to detect those traffic sign candidates simple circular and triangle shapes are used - see Fig. 5. A scale and rotation transforms are applied to those shape images in order to build templates of various sizes.

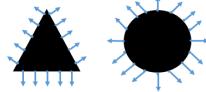


Fig. 5: The samples of warning and prohibitory sign template. Features (quantized gradient orientations) are superposed. In practice 100 features per template is used.

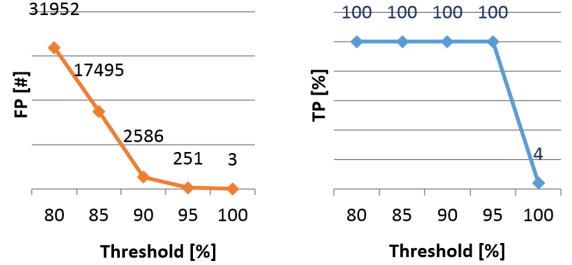


Fig. 6: The performance of the template matching procedure. The left graph shows the number of false positives in relation to the threshold - the minimum percentage of matched template features. The right graph represents the true positive percentage in relation to the threshold.

### C. Object tracking

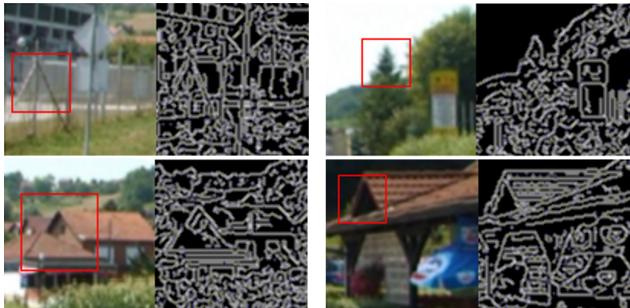
Object tracking is achieved using Kanade-Lucas tracker [11]. Some features due to extensive noise become invalid in a sense they're no longer matched correctly. This problem manifests in inaccurate tracked point speed and orientation. Those features are filtered out by using their magnitude and orientation as parameters as follows: mean and variance are calculated. All features which magnitude and orientation variance are within  $\pm\sigma$  are taken. Another difficulty is that objects that are supposed to be tracked are moving with the same speed as well as background so the additional filtering is hard to apply. The minimum number of detections per trajectory is set in advance. This phase greatly decreases the number of false positives. The second stage provides additional filtering using the border-redness feature Sec. III-D. Note that using temporal constraint a traffic sign may not be detected at a long range because those types of verifications are carried over several video frames.

### D. Color verification

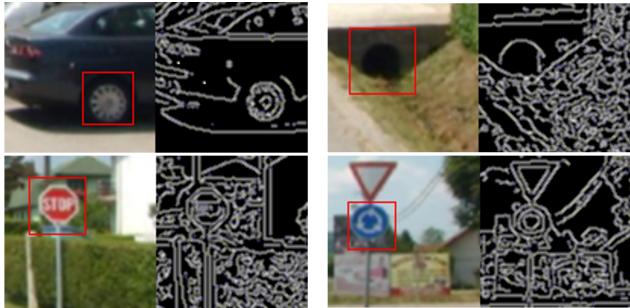
As shown in Fig. 7 false positives are still present after enforcing temporal constraint. The proposed verification step takes advantage of the template representation and traffic sign border redness - see Fig. 8 for details. The redness measure is computed as:

$$I_r = I_R - I_G \quad (1)$$

where  $I_r$  is redness image, and  $I_R$  and  $I_G$  are red and green image channels. Extracting just red image channel is not enough because many green objects have high red component, therefore the Eq. 1 is used to calculate image redness. After the initial match the best matched template is taken and normals are constructed at the feature locations - see Fig. 8. The length



(a)



(b)

Fig. 7: Samples of false positive detections that are remained after temporal consistency enforcement. a) Warning sign false positives b) Prohibitory sign false positives.

of the normal vector is proportionally resized according to the template scale as shown in Fig. 9. As the matched templates sometimes do not fit to the object's edge, the worst case is taken, where the matched templates are partly matched with both sign borders as shown in Fig. 10.

Therefore a length of normals positioned in each template feature position are increased in both directions. Pixels values are taken along normals where the bilinear interpolation is used. The obtained pixel values are smoothed in a way that pixel values are clustered by kNN [12] algorithm in two classes and a median operation is applied to the clustered array in order to obtain final class indices. The pre-processed array can have two minima corresponding to the inner and outer sign border or just one minima corresponding to the outer sign border due to low picture quality and small size where the redness spreads into traffic sign interior. There should be at least one transition from maxima to minima, jump, in order to classify the normal as valid. This procedure is taken for every normal as shown in Fig. 9. At the end valid lines are counted and compared against threshold.

#### IV. EXPERIMENTAL RESULTS

Validation of the proposed method has been conducted using the following parameter values. The number of features per template is set to 100, the neighbourhood size is 5 pixels and the maximum similarity between angles is set to 4. The number of quantized orientations (1st and 2nd quadrant) is 8. There are just two template prototypes: filled circle and filled triangle that represent the prohibitory and the warning

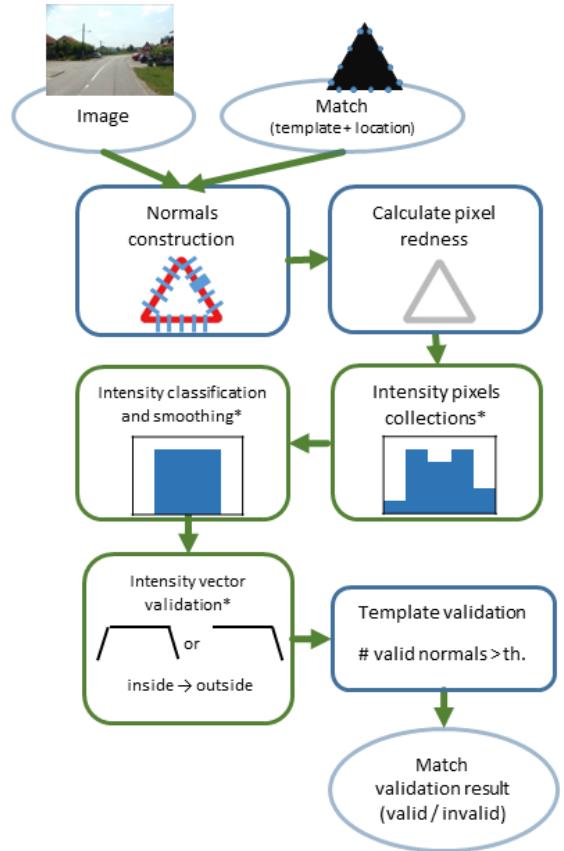


Fig. 8: Flow diagram of the color verification method. Operations executed on each normal are marked with \*.



Fig. 9: Color (redness) verification of a warning sign and a circular shape. Each image group shows original image, detected shape with superposed normals and image redness. Valid normals are marked with blue, red otherwise.



Fig. 10: A best template in a match area partly matches outside and inside circle in prohibitory sign or triangle is partly-matched in warning signs.

	Video 0	Video 1	Video 2	Video 3	Video 4	Video 5
Area (urban / rural)	rural	rural	urban / rural	urban / rural	rural	urban / rural
Duration (hh:mm:ss)	00:01:42	00:12:40	0:11:06	00:14:34	00:07:22	00:02:13
Prohibitory #	1	14	6	9	1	22
Warning #	3	12	17	17	5	2

TABLE I: Video sequences used for traffic sign detection and the number of signs in each sequence. *Video 0* is used for parameter settings.

signs. The original size of a template is 120x120. They are scaled until their size does not drop below 20x20 and rotated  $\pm 5$  degrees. There are 216 binarized templates (108 circles + 108 triangles), which are smoothed with Gaussian of  $\sigma$  value 3. The minimum number of matched features is set to 95% meaning that in this case 95 features must be correctly matched to mark the location as a candidate. The maximum number of concurrent verifications is 4. Each tracker can use up to 15 KL features. If there are less than 2 features an object is considered lost. There should be at least one detection in 3 frames and at least one positive verification. The minimum number of valid normal vectors in shape verification is 60 (60 % as 100 features per template is used).

A single video has been used for parameter selection during the training process. Table I shows information about the video sequences used in the experimental validation. The proposed method is evaluated on five very challenging videos with the length of more than 45 minutes in total. By using only template matching algorithm with a fixed threshold of 95 a very large number of false positive detections is obtained as shown in Fig. 7 and Fig. 6. The most important parameter is the template matching threshold, which is set to relatively high due to extensive background clutter in non-urban scenes. The performance of a detection without and with the post processing stage is shown in Table II.

## V. CONCLUSION

This paper presents a new method for detection of prohibitory and warning signs. The experimental validation has shown that the proposed traffic sign detection procedure can detect traffic signs in various conditions. By enforcing temporal consistency the number of false positives is significantly reduced, but without reducing the true positive rate. The color verification method in post-processing stage exploits pre-built templates and image redness as described thus reducing false positive rate even further. Although detection of only two traffic sign types is addressed the proposed method could be extended to detect other traffic signs by incorporating other templates into the existing set. In addition the presented method can be combined with machine-learning techniques to increase detection speed (e.g. fast template matching + Viola-Jones) or to efficiently reduce false positives (e.g. Viola-Jones + fast template matching).

## ACKNOWLEDGMENT

This research has been partially supported by the European Union from the European Regional Development Fund by the project IPA2007/HR/16IPO/001-040514 VISTA - Computer Vision Innovations for Safe Traffic.

## REFERENCES

- [1] S. Maldonado-Bascon, S. Lafuente-Arroyo, P. Gil-Jimenez, H. Gomez-Moreno, and F. López-Ferreras, "Road-sign detection and recognition based on support vector machines," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 8, no. 2, pp. 264–278, 2007.
- [2] C. Bahlmann, Y. Zhu, V. Ramesh, M. Pellkofer, and T. Koehler, "A system for traffic sign detection, tracking, and recognition using color, shape, and motion information," in *Intelligent Vehicles Symposium, 2005. Proceedings*. IEEE, 2005, pp. 255–260.
- [3] S. Varan, S. Singh, R. Sanjeev Kunte, S. Sudhaker, and B. Philip, "A road traffic signal recognition system based on template matching employing tree classifier," in *Conference on Computational Intelligence and Multimedia Applications, 2007. International Conference on*, vol. 3. IEEE, 2007, pp. 360–365.
- [4] M. Bénallal and J. Meunier, "Real-time color segmentation of road signs," in *Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on*, vol. 3. IEEE, 2003, pp. 1823–1826.
- [5] J. F. Khan, S. M. Bhuiyan, and R. R. Adhami, "Image segmentation and shape analysis for road-sign detection," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 12, no. 1, pp. 83–96, 2011.
- [6] X. W. Gao, L. Podladchikova, D. Shaposhnikov, K. Hong, and N. Shevtsova, "Recognition of traffic signs based on their colour and shape features extracted using human vision models," *Journal of Visual Communication and Image Representation*, vol. 17, no. 4, pp. 675–685, 2006.
- [7] M. Á. García-Garrido, M. Á. Sotelo, and E. Martín-Gorostiza, "Fast road sign detection using hough transform for assisted driving of road vehicles," in *Computer Aided Systems Theory-EUROCAST 2005*. Springer, 2005, pp. 543–548.
- [8] D. Gavrila, "Traffic sign recognition revisited," in *Mustererkennung 1999*. Springer, 1999, pp. 86–93.
- [9] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, no. 2, pp. 137–154, 2004.
- [10] S. Hinterstoisser, C. Cagniart, S. Ilic, P. Sturm, N. Navab, P. Fua, and V. Lepetit, "Gradient response maps for real-time detection of textureless objects," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 5, pp. 876–888, 2012.
- [11] J.-Y. Bouguet, "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm," *Intel Corporation*, vol. 2, p. 3, 2001.
- [12] N. S. Altman, "An introduction to kernel and nearest-neighbor non-parametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.

Sign type / Video	1	2	3	4	5	1	2	3	4	5
<b>Prohibitory #</b>	<b>14</b>	<b>6</b>	<b>9</b>	<b>1</b>	<b>22</b>	<b>14</b>	<b>6</b>	<b>9</b>	<b>1</b>	<b>22</b>
Detected	14	6	9	1	22	14	6	9	1	22
False discovery	42	20	12	5	0	1	2	2	0	0
Missed	0	0	0	0	0	0	0	1	0	1
<b>Warning #</b>	<b>12</b>	<b>17</b>	<b>17</b>	<b>5</b>	<b>2</b>	<b>12</b>	<b>17</b>	<b>17</b>	<b>5</b>	<b>2</b>
Detected	12	15	17	4	2	12	17	17	5	2
False discovery	40	45	30	51	0	1	2	1	2	0
Missed	0	2	0	1	0	1	2	0	0	0

a)

b)

TABLE II: Number of true positives, false positives and false negatives per each sequence using a) template matching algorithm with temporal consistency enforcement, b) template matching algorithm + temporal consistency enforcement + color verification. The number of false discoveries is obtained by counting how many false objects are detected in each video, thus it does not match the sum of true positives and false negatives.

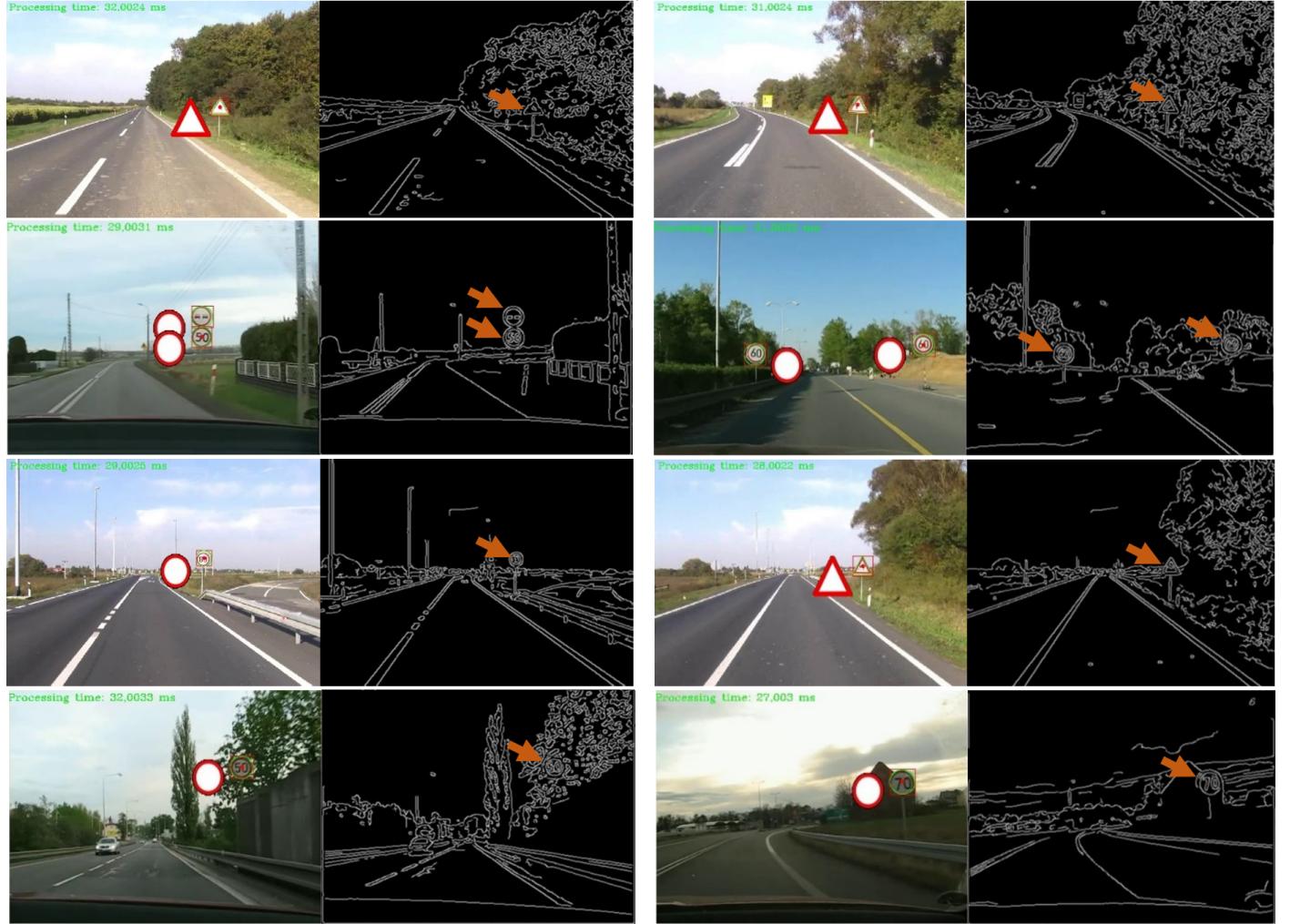


Fig. 11: Warning and prohibition sign detection. Each pair represents the original image with found detections and the edge magnitude image. The obtained results from template matching procedure are filtered by traffic sign candidate tracking and color border verification as described in III-C and in III-D respectively. The execution time is shown in the upper-left corner of the each original image. All tests are conducted on Intel Core i5-2500K CPU.