A Fast and Robust Traffic Sign Recognition

Ali BEHLOUL and Yassmina SAADNA
LaSTIC Laboratory at the University Hadj Lakhdar, Batna, Algeria

ABSTRACT: Traffic Sign Recognition (TSR) system is an important component for the intelligent vehicles, it can assist and inform the driver about dangerous situations such as stop, icy roads, no entry or speed limits. In this paper we present a fast and robust traffic sign recognition system constituted of three modules which are: segmentation, detection and recognition of sign type. In the first module we start by applying a filter after normalization of the three RGB channels to extract red, green, blue and yellow maps. To detect the signs and identify their forms, in the second module we propose a new and fast approach for pattern recognition based on minimum bounding rectangle. For the third module, the recognition is made by using a matching directly between the SURF descriptors of the detected traffic sign and the traffic signs in the database, in this module we apply a filtering interest points detected and we keep only the points that are inside the pictogram’s sign. The evaluation of the proposed approach gives good results compared to some powerful techniques. As a result, with the proposed system we have obtained a high performance with 95.65% sign detection, 97.72% traffic sign identification and 89.59% traffic sign recognition rate in an average time less than 80 ms/image.

KEYWORDS: Advanced driver-assistance systems, pattern recognition, image segmentation, road traffic, interest points, image color analysis.

1. INTRODUCTION

Automatic traffic sign detection and recognition, as an important task of Advanced Driver Assistance Systems (ADAS), has been of great interest in recent years. Traffic sign detection has a direct impact on the safety of driver, and damages can be easily produced due to their ignorance. Automatic systems developed to assist the driver [1], [2], [3], based on detection and recognition of signs can consequently correct the most unsafe driving behaviors.

To identify the signs, the most part of researchers divide the task in a three sequential stages: segmentation, detection and recognition. The role of the first stage is to determine the area of the signs in the road scene, whereas the detection selects areas that have an appropriate traffic sign shape. Next, recognition stage identifies the information of the extracted traffic signs. The approaches of identifying the signs can be classified into three main classes: geometrical methods which are based on the geometric shape to detect signs [4], [5], colorimetric methods which are based on the color in the detection phase [6], [7], [8], [9], and there are also methods that combine learning with the methods in the previous two classes [10], [11], [12].

In colorimetric methods the most popular color spaces used are RGB, HSV, HSI, YCbCr, and CIELab. In these approaches images are first converted to a designated color space and then a segmentation algorithm based on dominant color detects interest regions, these regions are then filtered by a recognition algorithm or model appearance. Escalera et al. [8], chose the HSI space to detect the signs, they considered only the H and S components to overcome the problem of brightness variations and use a thresholding to determine the dominant color to segment the image. RGB color space is used to threshold the image in [9]. Piccioli et al. [6], use color and a priori information to determine the interest regions and limit the possible locations of signs in the image. In another approach proposed by Chen et al. [13], based on SVF (Simple Vector Filter)
which was proposed in [14], they found in their experiments that the approach gives good results in the separation of elements of red, blue and yellow.

In the case of geometrical methods, the detection of signs is made from image contours which are analyzed by a structural or global approach. These methods are generally more robust than colorimetric because they deal with the gradient of the image, and can process images in grayscale. A Hough transform was used by [5], to detect edges of the signs, after they select only the closed contours; recognition of candidates is done by a neural network. In [8], authors chose to use another technique called Radial Symmetry Transform (TSR) to detect speed limit signs, also it was used by [15]. Paulo et al [7], used a fast radial symmetry transform (FRST) detection method to identify circular shapes and Harris corner detection algorithm for triangular and square shapes. The authors in [16], detected circular and triangular signs using Color Distance Transform (CDT), the distance is computed for each color channel separately. To recognize the shape of the sign, the authors in [17], used the vector of Distance to Border (DTB), which is the distance between the outer contour and its bounding rectangle for each segmented blob there are four vectors DTB (left, right, top, bottom). In [4], authors developed a colored traffic sign recognition system based on Scale Invariant Feature Transform (SIFT).

In the approach [10], a cascade detectors increasing complexity is used, each detector is a set of classifiers based on Haar wavelet, these classifiers use a learning algorithm AdaBoost, a threshold in the HSI color space is used for segmenting a candidate blobs, the blobs are classified according to their shape by a linear support vector machine (SVM), and the classified blobs are recognized according to their patterns using non-linear SVM [18]. In [12], a threshold is applied over a HSV color space to segment the image, to classify the shape of an extracted candidate a method based on support vector machines (SVM) with a Gaussian kernel is used. To detect sign an evolutionary Adaboost and a classification through a Forest Error Correcting Output Codes (F-ECOC) framework are used in [11]. The authors in [19], proposed for recognition traffic sign to combine knowledge-based analysis and radial basis function neural classifier (RBFNN).

The rest of this paper is organized as follows: the next three sections present the proposed traffic sign detection and identification system. Section 2 presents the color segmentation of a traffic sign with a filtering method. In Section 3 we discuss the proposed approach to detect and identify signs in the segmented maps. The recognition approach of the sign type is described in Section 4. Section 5 illustrates the experimental results. Finally, the conclusions are given in Section 6.

2 TRAFFIC SIGN SEGMENTATION

The color information is a discriminative feature road signs, there are red, blue, and yellow signs, and to extract it a segmentation of image is required. Many authors use the RGB color space to avoid the time of conversion to another color space.

Chen et al [13], use the Simple Vector Filter (SVF) proposed by Asakura [14], they found in their experiments that this approach gives good results in the separation of elements in red, blue and yellow. For each pixel \( x = [X_R, X_G, X_B] \) SVF is calculated using the following equation:

\[
\begin{align*}
\text{Red: } & X_R - X_G > 40 \text{ and } X_R - X_B > 40; \text{ else 0; } \\
\text{Blue: } & X_B - X_R > 65 \text{ and } X_B > 45; \text{ else 0; } \\
\text{yellow: } & X_R - X_B > 50 \text{ and } X_G > 50; \text{ else 0; }
\end{align*}
\]

In this approach the rate of unwanted pixels is very small compared to previous approaches, but this approach is closely linked to thresholds, in some pictures where there is a great change in lighting, signs are not detected.

In another approach [16], a filter is applied for each pixel of the RGB image by a set of transformations:

\[
\begin{align*}
f_R(x) &= \max(0, \min(x_R - x_G, x_R - x_B)) / S \\
f_G(x) &= \max(0, \min(x_B - x_R, x_B - x_G)) / S \\
f_Y(x) &= \max(0, \min(x_R - x_B, x_G - x_B)) / S
\end{align*}
\]

\[
S = x_R + x_G + x_B
\]

With \( S = x_R + x_G + x_B \)

In this approach signs have a high intensity, which facilitates their detection after thresholding of the image, but it also generates unwanted noise pixels in the yellow map such the green color is seen as yellow.

King et al. [19], normalize the three RGB channels by the intensity I to avoid the problem of lighting, where:

\[
I = \frac{R + G + B}{3}
\]
and $r, g, b$ are the normalized red, green and blue channels of the input image.

$$\begin{align*}
r &= \frac{r}{T}, \quad g = \frac{g}{T}, \quad b = \frac{b}{T}
\end{align*}$$

(4)

The maps red, green, blue and yellow are constructed by the following equations:

$$\begin{align*}
R' &= r - \frac{g + b}{2} \\
G' &= g - \frac{r + b}{2} \\
B' &= b - \frac{r + g}{2} \\
Y' &= \frac{r + g}{2} - \frac{|r - g|}{2} - b
\end{align*}$$

(5)

Signs with this approach are distinguished, but in the results of our experiments when comparing the red map with the same map of the previous method [13], we found in this approach that the rate of unwanted pixels are high, and unwanted pixels present in the red map are also present in the green or yellow maps.

In our approach we have improved the approach proposed in [19], the idea of our filter is to remove unwanted pixels in the red and yellow maps.

The filtering procedure is described as follows:

1. A pixel $p(x, y)$ which has a high intensity in the red map and green or yellow maps, this is an undesirable pixel in the red map, but not necessarily in the yellow map.
2. To be undesirable in the yellow map we use a threshold, the pixel is considered undesirable if its blue or green value is below the threshold $\alpha$, the value obtained after experiments is: $\alpha = 180$.

Fig.1 present the results of the proposed filter, it shows a comparison between the images before and after filtering. If we see the results in Fig.1 (e, f, g) of our filtering we found that we have eliminated almost unwanted pixels and we obtained the best results in the yellow map.

Fig. 1. Effect of the proposed filtering, (a) original image RGB, (b) (c) (d) segmented images before filtering; (e) (f) (g) after filtering segmented images

Red map
Blue map
Yellow map
3 TRAFFIC SIGN DETECTION

After the extraction of pattern maps with the proposed filtering method, the signs are located and at this stage it is necessary to distinguish them from their backgrounds to recognize their forms. This distinction is made by detecting limits of the sign based on its contour. We have proposed a new fast approach to pattern recognition, in this approach we used a rectangle that encompasses the detected contour to characterize its shape. To detect the shape of a sign we used a score of intersection between the detected pattern and the four lines of the rectangle.

3.1 CALCULATION OF THE MINIMUM BOUNDING RECTANGLE

To recognize the shape of the sign candidate, the idea of our approach is to compare the detected pattern with the rectangle that encompasses it named “BoxOut”, we calculate a score of intersections between the contour of the pattern and the four lines of the BoxOut as follows:

- Score0: If there is no intersection between the contour and the line of the BoxOut;
- Score1: If there is a small intersection in few points between the contour and the line of the BoxOut;
- Score2: If the intersection occupies nearly a quarter of the line of the BoxOut;
- Score3: If the intersection occupies almost all the line of the BoxOut.

In TSR we have the advantage that there are four forms: rectangle, triangle, circle and octagon, which facilitates work and helps us to deduce the type of any sign based on the score of intersection calculated. Fig. 2 shows the proposed approach.

![Fig. 2. The proposed detection approach (a) rectangle, (b) triangle, (c) circle, (d) octagon](image)

By observing Fig.2 we find that each form has a different distribution of scores, which allows us to consider as a discriminant characteristic:

- Rectangle is composed of (score3, score3, score3, score3);
- Triangle is composed of (score3, score0, score1, score0);
- Circle is composed of (score1, score1, score1, score1);
- Octagon is composed of (score2, score2, score2, score2).

To ensure that the proposed approach is robust to rotation, we seek the minimum rectangle that can encompass the contour, as shown in Fig.3.
Fig. 3. Illustration of using the minimum rectangle that encompasses the form

For finding the minimum rectangle that encompasses the contour we use a specific structure to represent the rectangle, as shown in Fig.3 the information required for the representation are: the gravity center of the rectangle, the height, the width and the angle of rotation, which are obtained by using the cvMinAreaRect2 function of the OpenCV\(^1\) library, once we have the height, the width and the center of gravity it’s easy to calculate the coordinates of the rectangle.

As we explained above, to recognize the form we must calculate an intersection between the contour of this form and the BoxOut that encompasses it. Made that the four points of the BoxOut are obtained, they will be used to calculate the functions of each line, and the calculation is provided by a:

\[
\Delta_1 = a_1x + b_1 \\
a_1 = (P_2.y - P_1.y)/(P_2.x - P_1.x) \\
b_1 = P_1.y - a_1 * P_1.x
\]  

(6)

Fig. 4. Illustration of the four lines of the BoxOut

The same for all other lines (\(\Delta_2, \Delta_3, \Delta_4\)) of the BoxOut. And to reduce the noise problem we proposed to calculate for each line \(\Delta\) two parallel lines \(\Delta'\) \(\Delta''\), the whole forming a band instead of a straight line such that:

\[
\Delta': y = ax + (b - \varepsilon) \\
\Delta'': y = ax + (b + \varepsilon)
\]  

(7)

\(^1\) http://opencv.org/
Now the score of intersection is calculated between the outline of the shape and the four bands of the BoxOut.

The Bilateral Chinese Transform (BCT) proposed by Belaroussi et al. [20], detects the circular and polygonal signs, but to detect the triangular signs they combine between the detection of the peaks and the center of the triangle, and for the rectangular signs they did not specify what type of transformation used. The major disadvantage of this approach is that for each type of sign they use a different transformation, they use three detection algorithms ellipse, triangle and quadrilateral, the final form will be chosen depending on the degree of compatibility provided by each of the algorithms. Against in our approach we use a single method to recognize the form of the candidate, so one treatment is performed instead of three treatments.

The Radial Symmetry Transformation in (TSR) is used in [8], to detect speed limit signs, but the disadvantage of this approach for triangles signs is that it can only find the size and position of the forms: it cannot distinguish a give way sign and intersection sign. Whereas our approach can distinguish between them because we know on which line of the BoxOut there is an intersection with the contour.

The approach proposed by Reina et al. [21], applies a reorientation on the detected signs to make it robust to rotations, but our approach is robust to rotation with no need to reorient the form. The transformation distance used by Ruta et al. [16], to solve the problem of scaling requires several models, one model for each shape on different scales, even with the use of a hierarchy of models their approach is very time consuming process. While our approach is fast and robust to changes in scale as it calculates a score of intersection just compared to the width or height of the BoxOut.

The vector of the DTB (Distance to Borders) proposed in [17], their approach is robust to the rotations and translations but not really to changes in scale because they zoom all candidates signs to 36×36. Thus the major drawback is the size of DTB vector, for each candidate there is four vectors DTB of size 1×36, but in our approach we use four scores of intersection, the global vector will be of size 1×4, which proves its robustness to scale changes.

4 TRAFFIC SIGN RECOGNITION

The color and shape of the sign are identified at this stage, to recognize the sign type of a query image we compare it with a database of signs, which are classified according to their color and shape in order to accelerate research. The comparison is based on the calculation of a matching measure to determine the nearest sign.

To use a matching directly by calculating the correlation of descriptor between the query image and the images in the database is highly sensitive to the background noise present in the query image, that is why we chose to use a matching based on interest points for their robustness to noise, rotation and illumination changes. In our work we have chosen to use the SURF descriptor because the comparative study of Juan et al. [22], demonstrates the superiority of SURF descriptor against SIFT and PCA-SIFT in terms of the runtime performance and robustness to illumination changes.

![Fig. 5. The signs models, (a) indication, (b) warning (c) Prohibition (d) obligation](image-url)
In this phase we try to match pictograms signs, we apply filtering to the interest points detected in the query images to delete points on the background and keep only the points that are inside the pictogram, we have created for each class of signs a sign model to filter the interest points detected as shown in Fig.5. The interest points detected outside the black sign model are rejected, in order to leave the interest points that represent the pictogram.

5 The matching

The matching is used to reduce the semantic gap between low-level (extraction and description) and high-level (recognition) while seeking the pair of points with a high similarity, each point of the query image is associated to the nearest interest point in image of the database.

Marius in [23], uses the principle of random KD trees proposed in [24]. KD trees are used to structure the search space to accelerate the comparison of an element with the others, but the search performance of a KD trees are close to those of a linear search when the size of the data space is large. KD trees would not be as effective for SURF with 64 dimensions. As the number of interest points does not exceed twenty points and SURF descriptor size is 64, the Brute Force algorithm matcher appears efficient.

Once we found the right match between the interest points of the pictogram and those of the image of the database, it remains to measure this matching to select the closest image. Computing a measure of correlation between the descriptors of matched points of the two images, then to choose the pair having a maximum correlation does not give good results, for this we chose to combine a coefficient $\sigma_{pq}$ with the sum of distances as follows:

$$Dist(p,q) = \frac{\sum_{j=1}^{N} \text{dist}(p_j, q_j)}{\sigma_{pq}}$$

With $\sigma_{pq} =$ number of good matches

The number of good matching between the descriptors of two images is calculated by setting a maximum distance between the descriptors, that we keep the pairs of descriptors that have a distance below the threshold. The closest image is the one that has a minimum distance $Dist$, an example of matching results is shown in Fig.6.

6 Experimental results

Our system is divided in to three modules: the segmentation module, the signs detection module and the signs identification module. The segmentation module is used for locating the sign it generates four maps after segmentation, the detection module uses maps to detect the sign and recognize its geometric form, in the identification module we detect interest points in pictogram sign to identify it.
Our traffic sign recognition system has been developed based on visual C++ using the Intel® Open Source Computer Vision Library (OpenCV). To evaluate our system, experiments were performed on the real data. The database traffic sign image consists of 48 images with 360×270 pixels containing three different traffic signs\(^2\).

To evaluate the performance of our system we evaluated each module separately, we use a series of measures, we divide all the tests into two groups positive examples and negative examples, and we define:

- **TP**: the number of true positives, the number of true signs detected or identified;
- **FP**: the number of false positives, the number of non-detected or identified signs;
- **TN**: the number of true negatives is the number of non-signs undetected;
- **FN**: the number of false negatives is the number of true signs undetected.

### 6.1 Evaluation of Segmentation Module

The segmentation module is very important because it consists of locating and detecting the maximum road signs and rejecting most of the other objects. The evaluation of the segmentation module consists in measuring the capacity of the module to locate the signs and ignore the background of the scene, which facilitates the detection of signs.

Our segmentation module is able to segment correctly 46 of 48 road scene images with only 1 false positive in an average time with filtering less than 7 ms/image. We note that the proposed filter gives good results, where the rate of FP eliminating in the maps generated with the approach [19], achieved 85.71% and 90% with maps generated with the approach [16]. Table 1 summarizes a comparison of results.

![Table 1. Performances of the segmentation module.](image)

<table>
<thead>
<tr>
<th>Category methods</th>
<th>Number of images</th>
<th>TP (signs)</th>
<th>FP (signs)</th>
<th>FN (signs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our method</td>
<td>48</td>
<td>46</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Method [19]</td>
<td></td>
<td>46</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Method [16]</td>
<td></td>
<td>44</td>
<td>10</td>
<td>4</td>
</tr>
</tbody>
</table>

We got very good results of segmentation; we can well segment 46 of 48 road scene images with 1 false positive and 2 false negative.

### 6.2 Evaluation of the Detection Module

This module consists of detecting road signs and estimates their geometry to properly identify the class of candidate sign, in this module we try to detect the shape of signs in the maps generated by the segmentation module. Of 46 images provided by the segmentation module, the traffic signs were extracted correctly 44 images. In table 2, the sign detection module achieved 95.65% of correct detection rate. In addition to the high detection rate, the detection time does not exceed 1 ms/image. The number of FP in this phase is the FP results generated by the segmentation phase.

The high rate of detection demonstrates the ability of the proposed approach to detect signs, our approach arrive to detect signs in images even in different condition (lighting, rotation and zoom).

---

\(^1\) The image database of traffic signs is available at: [http://www.cs.rug.nl/~imaging](http://www.cs.rug.nl/~imaging)
Table 2. Performance of the detection method of geometric signs shape.

<table>
<thead>
<tr>
<th>Category</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>46</td>
<td>44</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

The final evaluation part is dedicated to evaluating the recognition module of signs.

6.3 Evaluation of the Recognition Module

This module is to recognize the sign detected by matching it with the signs of the corresponding class. The matching is applied on the SURF descriptors computed around interest points detected on the candidate image and the images of the database.

Before filtering interest points we identify from 44 signs detected in the detection phase 31 with 2 false positives, but after filtering, where we keep only the interest points representing the pictogram, we correctly identify 43 from 44 signs with only 1 false positive because we set a threshold on the measure shown in Equation (8) and we accept only the signs with a distance less than 0.70; from 44 images candidates the recognition module achieved 97.72% of correct recognition rate. Table 3 shows the performances of the recognition module. With filtering interest points we can identify 43 candidate images of 44 images with only 1 false positive.

Table 3. Performances of the recognition module.

<table>
<thead>
<tr>
<th>Filter</th>
<th>Number of targets</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before filtering</td>
<td>44</td>
<td>31</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>After filtering</td>
<td>44</td>
<td>43</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4 summarizes the evaluation of our system, we present the rate of TP, FP and FN for the different modules and the entire system. In the segmentation phase we detect signs with 95.83% and 2.08% of FP, and in detection phase 95.65% of the segmentation results are correctly detected with 2.17% of FP, in the recognition phase 97.72% of the signs detected are correctly identified with 2.27% of FP.

Table 4. Performances of our traffic signs recognition system

<table>
<thead>
<tr>
<th>Module</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segmentation</td>
<td>95.83%</td>
<td>2.08%</td>
<td>4.16%</td>
</tr>
<tr>
<td>Detection</td>
<td>95.65%</td>
<td>2.17%</td>
<td>4.34%</td>
</tr>
<tr>
<td>Recognition</td>
<td>97.72%</td>
<td>2.27%</td>
<td>2.27%</td>
</tr>
<tr>
<td>System</td>
<td>89.59%</td>
<td>6.25%</td>
<td>10.41%</td>
</tr>
</tbody>
</table>

The results obtained by our system are very encouraging, we are able to correctly identify 89.59% of the images processed by our system in an average time less than 80 ms/image. The proposed approach not only provides accurate identification signs, but it is insensitive to the differences appearance of the signs (lighting, rotation and zoom) in the real world.

Some detection simulation results are shown in Fig.7. Traffic signs can be detected and identified correctly in various color and shape.
7 Conclusion

In this paper we have presented a system for detection and identification of traffic signs in a color image. Our system is divided into three phases: The first one concerns the segmentation and location of signs in images, the second is for the detection of the geometric shape of signs already located and the third is to identify the type of signs. In the first phase the filtering applied on the generated maps eliminates almost all unwanted pixels. In the phase of detection form we have proposed a method for pattern recognition by using a minimum bounding rectangle that encompasses the contour, this method detects the shape of the sign with a high detection rate of 95.65%. The detected signs are extracted and passed to the recognition phase. To identify the type of sign we applied a matching between descriptors SURF of sign detected and those of images in the database signs. In this phase we obtained a high identification rate of 97.72%. As a result, the proposed solutions have allowed developing a system with high performance with 89.59% identification and recognition rate. In this paper only single images are considered. In future works, we will focus on the implementation of robust traffic sign in video traffic sequences.

References


