Morphological Classification for Traffic Sign Recognition

Hai Nguyen Thanh

Faculty of Electrical and Electronics Engineering, HCMC University of Technical Education, Vietnam

Abstract In this paper, a novel method is proposed for the Traffic Sign Recognition (TSR) using the Principle Component Analysis (PCA) and the Multi-Layer Perceptron (MLPs) network. In particular to the proposed morphological classification method, the candidate signs are individually detected from two components of the YCbCr space and then classified into three shape classes: circle, square, and triangle based on computing the rotated version correlations. The PCA-based features of these sign objects will be used for the MLPs as the training system corresponding to previously determined class. This approach not only reduces the time but also increases the performance in the recognition process. In simulation, the proposed method is estimated with over 500 statistic images and its accuracy rate is up to 96%.

Keywords Traffic Sign Recognition, Principle Component Analysis, Multi-Layer Perceptron, Morphology Classification

1. Introduction

Intelligent Transport System (ITS) has become the great solution for drivers in protection and improvement of life in recent decades. As one of the important fields in the ITS, the traffic sign recognition system immediately represents the current traffic situation to give the warning from road signs [1-2]. In addition, it suggests drivers some useful advices for safety and convenience. In this paper, the morphological classification for traffic sign recognition is proposed to alert traffic participation.

In feature extraction for recognition, Lafuente-Arroyo et al. [3] proposed an optimization in the traffic sign identification task on the spatial domain. It could be seen as the study of a Two-Dimension (2D) reduction method based on the Principle Component Analysis (PCA) for optimal recognition. In the letter [4], Deli Pei et al. extended the original unsupervised model with an additional supervised term to restrain the classification errors of the recovered feature representations, called the supervised low-rank matrix recovery model. As an effective method, the eigen-based Traffic Sign Recognition (TSR) [5] applied a PCA algorithm to extract the most significant components of the input images for categorization. A set of weights was computed from most effective eigenvectors of a database and then unknown objects would be classified using the Euclidean distances. The high accurate rate of this method is around 97%.

In another feature extraction of various traffic sign shapes, the process sequence to extract traffic signs including color segmentation, shape simplification, and shape decomposition was presented [6]. The recognition was achieved through the direct matching to templates for closed candidate shape and computation of the minimum geometric difference between object and template for unclosed shapes. A novel approach to detect and recognize traffic signs based on ridge regression [7], in which a precise segmentation in the RGB color space was obtained with the same performance as other learning machines. In order to resist the illumination variations and distortions, features were extracted using the Otsu method for a recognition process. However, this method is inappropriate for the subsequent frames.

Object recognition plays an important role in detecting traffic signs. One of recognition methods is that the Support Vector Machines (SVMs) for image detection and the Gaussian-kernel SVMs for content recognition have been utilized [8]. In detail, this SVM algorithm used the colour-based segmentation to identify all various shape signs such as, circular, rectangular, triangular and octagonal. Ke Lu et al. [9] proposed a novel graph structure which is a balance between local manifold structures and global discriminative information of traffic signs. The result is that this algorithm allows to recognize traffic signs with better performance than the previous methods. As the main component of the Driver Assistant Systems (DAS) [10], a real-time traffic detection and classification method for circular and triangular signs was proposed using the HOG-based SVM algorithm.

In order to improve the accuracy rate, the segmentation based on the enhancement of red color channel was utilized in an identification process using the tree classifiers. In addition, candidate regions as Maximally Stable External Regions (MSERs) [11] were automatically detected before they were recognized based on a cascade of the SVM
algorithm. Another method is that the TSR was introduced as the Advanced Driver Assistance System (ADAS) [12], in which the combination of Viola and Jones (VJ) object detection and the Hue, Saturation, Intensity (HSI) colour space were represented. The detection process in this scheme found an optimal set of candidates using Adaptive Boosting (AdaBoost) cascades. Therefore, a hierarchy of the SVMs was employed to control the recognition of traffic sign types.

A Multi-Column Deep Neural Network DNN (MCDNN) algorithm from 25 nets, 5 per pre-processing method was proposed to classify the German traffic sign with a recognition rate of 99.46%. Moreover, this method increases robustness to various types of noise [13]. Moreover, this network with 25 nets increased the recognition rate. In addition, a previous method for video detection of road signs based on computation of color eigenvectors was proposed [14], in which the candidate signs from road scenes were detected and a global model was built to detect and new road signs. Thus, all road sign candidates were recognized using the Radial Basis Function (RBF) network. Experimental results proved that this method for road signs detection is robust, powerful and accurate.

In this research, a novel TSR method using the morphological detection and classification is proposed to identify candidate objects for recognition. Based on the examination of correlations of each object and morphological samples in rotation, the decisions for the class of these objects are made. Therefore, the recognition using the combination of the PCA and the MLPs network becomes easier and faster with the high accuracy rate. This paper is organized as follows: the section II described the proposed morphological detection and classification. The recognition based on the PCA-MLPs network will be represented in section III. Finally, the experimental results and the conclusion are written in section IV and V in this paper, respectively.

2. Morphological Classification

Due to the chrome of traffic signs, the scheme, as shown in Fig. 1, for detection of doubt objects based on the Canny method for separated component in the YCbCr color space is proposed. The traffic signs are filtered by the scale condition to reject unsuitable objects and then they are classified into three classes by mapping their shapes to the samples.

![Figure 1. Block Diagram of the morphological detection and classification of traffic signs](image-url)

2.1. Doubt Region Extraction Using Contrast Enhancement

In this research, the effective detection method based on the morphological analysis is applied. The innovation of this method is that candidate objects are detected and categorized based on the morphology to reduce the time and to also improve the accuracy rate of the recognition process. Thus, the Canny segmentation is employed as the edge detection to find edges by looking for the local maxima of the gradient magnitude. In particular, the level of edge can be controlled through the threshold value. In order to improve a capability of detecting more doubt objects, especially, the RGB color images with more colors are converted and then separated into three components: the luminance Y, the blue-difference Cb, and the red-difference Cr. The YCbCr-based analysis can be considered as the effective solution for the traffic sign system with red and blue colors using for most of signs as shown in Fig. 2. In order to enhance sign objects, the Cb and Cr components have been adjusted using the following equation:

\[ g_0 = L \times \left( \frac{g_i - g_{\min}}{g_{\max} - g_{\min}} \right) \]  

where \( g_i \) and \( g_0 \) are the gray values of the input image and the output image \( g_{\max} \) and \( g_{\min} \) describe the maximum and minimum gray values of the input image, \( L \) is the range of the output histogram.
The area containing sign objects can be clearly observed when applying the enhancing algorithm as shown in Fig. 2b and Fig. 2e. In additions, the binary output images of the sample after segmentation with the Canny method are shown in Fig. 2c and Fig. 2f. In order to determine doubt regions containing signs, the holes, a set of background pixels, which can not be reached by filling on the background from edges, are blocked up in the binary images as shown in Fig. 3a and Fig. 3b. The fusion of them is represented in Fig. 3c using the following equation:

$$F = C_b \cup C_r$$  \hspace{1cm} (2)

where $C_b$ and $C_r$ denote the images after filling holes corresponding to two separated components.

**Figure 2.** The Canny segmentation for two separated color components. The first one with the images (a), (b), (c): Cb-component; the second with the images (d), (e), (f): Cr-component

**Figure 3.** Detection of doubt regions based on the fusion of two Cb and Cr components, (a) Cb-component, (b) Cr-component, (c) The fusion image
2.2. Morphological Classification Based on Correlation

These objects are individually extracted into the separated images and categorized based on their morphology. The main idea is comparing these objects with the samples corresponding to three classes such as triangle, circle, and square as shown in Figs. 4a - 4c. Classification of these objects, that can reduce the time of recognition on a neural network, can be listed in detail through the following steps:

2.2.1. Removal of Objects with Unsatisfactory Size

![Image of objects](image1.png)

Figure 4. Three samples corresponding to three groups for classification and the objects have been extracted as the individual images, in which (a) Triangle, (b) Circle, (c) Square, (d) 8 Doubt Objects

The bounding boxes of these objects as shown in Fig. 3c are usually the square (for circle or square signs) or the rectangle (for triangle signs). Therefore, the ratio of width \( C \), which has been set at the range \(0.8 \leq C \leq 1.2\), is written as follows:

\[
C = \frac{x_{\text{width}}}{y_{\text{width}}}
\]

(3)

where \( x_{\text{width}} \) and \( y_{\text{width}} \) denote the size of the extracted object, and \( C \) can be seen as the size condition in this research.

2.2.2. Computation of the Normalized Correlation Factors

The basic ideal is that the correlations of the doubt objects are compared with each sample for categorization. In particular, both sample and object can rotated with the angle \( \alpha = 30 \) degrees for determination of the correlation factors. The decision object class is calculated using the following average equation:

\[
m_{nc2,o,s} = \sum_{i=1}^{N} nc2_{o,s,i}
\]

(4)

where \( m_{nc2,o,s} \) is the mean of normalized correlation factors between the object, \( o \) and the sample, \( s \); \( nc2_{o,s,i} \) is the normalized correlation when rotating with the angle \( \alpha \) degrees for both, and \( N=360/\alpha \) is the total number of rotated versions. The equation for computing the correlation is expressed as follows:

\[
nc2 = \sum_{i=1}^{m} \sum_{j=1}^{n} O(i, j) \times S(i, j) \frac{1}{m \times n}
\]

(5)

where \( O(i, j) \) and \( S(i, j) \) are the gray values located at the coordinate \((i, j)\) of the object and the sample which have been normalized in the same mxn size (set at 80x80). It is noted that the values of \( O(i, j) \) or \( S(i, j) \) are the set 1 for bit 1 and -1 for bit 0. Therefore, the value of \((O(i, j) \times S(i, j))\) is either 1 or -1.

For each sign object, three values corresponding to the means of correlation factors between the object and three samples are determined. The sign object is chosen as the candidate sign for recognition by comparing these coefficients to the threshold referring the acceptable level of the object. In particular, the algorithm for classification can be described based on the consideration of the maximum means:

\[
\text{max}_o nc2 = \text{max}(m_{nc2,o,s=\text{circ}}, m_{nc2,o,s=\text{square}}, m_{nc2,o,s=\text{triangle}})
\]

(6)

<table>
<thead>
<tr>
<th>Case</th>
<th>Value ( C )</th>
<th>Threshold ( y )</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UF</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>Smaller</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>Larger</td>
<td>( m_{nc2,circ} )</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>Larger</td>
<td>( m_{nc2,square} )</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>Larger</td>
<td>( m_{nc2,triangle} )</td>
</tr>
</tbody>
</table>

Table 1. Conditions for classification (F: Fit, UF: Unfit)

<table>
<thead>
<tr>
<th>Object</th>
<th>( C )</th>
<th>( m_{nc2,circ} )</th>
<th>( m_{nc2,square} )</th>
<th>( m_{nc2,triangle} )</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.659</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Reject</td>
</tr>
<tr>
<td>2</td>
<td>1.552</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Reject</td>
</tr>
<tr>
<td>3</td>
<td>0.641</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Reject</td>
</tr>
<tr>
<td>4</td>
<td>0.957</td>
<td>0.275</td>
<td>0.274</td>
<td>0.429</td>
<td>Reject</td>
</tr>
<tr>
<td>5</td>
<td>1.529</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Reject</td>
</tr>
<tr>
<td>6</td>
<td>1.351</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Reject</td>
</tr>
<tr>
<td>7</td>
<td>0.949</td>
<td>0.851</td>
<td>0.567</td>
<td>0.217</td>
<td>Circle</td>
</tr>
<tr>
<td>8</td>
<td>1.029</td>
<td>0.817</td>
<td>0.675</td>
<td>0.249</td>
<td>Circle</td>
</tr>
</tbody>
</table>

Table 2. Results of the classification for a sample with \( y = 0.8 \)

Categorization is produced based on comparing \( \text{max}_o nc2 \) with the threshold \( y \) as shown in the Table 1. The result of classification for the sample image is represented in the Table 2, in which only two objects 7 and 8 detected as the candidate signs, are recognized in the circle sign database. Based on the positions, two sign objects are extracted as the individual images to normalize as shown in Fig. 5.
3. Traffic Sign Recognition Using PCA-MLPs Algorithm

For the recognition problem in this paper, the authors used the PCA algorithm to extract features traffic sign image which then can be propagated on the MLPs network with the previous trained set. In order to be more details, the recognition algorithm is described in Fig. 6.

3.1. PCA Algorithm for Sign Images

In recent years, the Principal Component Analysis (PCA) that alters the data into the new structure based on its variance has been used as the solution for face recognition [9] and others. In particular, a 2-D image is represented as a 1-D vector by concatenating each column (or row) into the vector as follows:

\[
X_j = \begin{bmatrix}
x_{1,j} \\
x_{2,j} \\
\vdots \\
x_{N,j}
\end{bmatrix}
\]

where \( N \) is the total pixels in an image, the \( x_{1,j}, x_{2,j}, \ldots, x_{N,j} \) are the gray values of the \( j^{th} \) image.

For \( M \) images with the same size, the set of the 1-D vectors is described using the following equation:

\[
X = \begin{bmatrix} X_1 & X_2 & \cdots & X_M \end{bmatrix}
\]

where \( X_1, X_2, \ldots, X_M \) are the 1-D vectors of the images.

Let \( m \) represent the mean image, one has:

\[
m = \frac{1}{M} \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,M} \\
\vdots & \ddots & \vdots \\
x_{N,1} & x_{N,2} & \cdots & x_{N,M} \end{bmatrix} = \begin{bmatrix} m_1 \\
\vdots \\
m_M \end{bmatrix}
\]

where \( m_1, \ldots, m_M \) are the images.

The images \( X_j \) centered by subtracting the mean image from each image vector are represented using the following equation:

\[
w_j = X_j - m
\]

where \( w_j \) is the set of 1-D vectors.

From Eq. (9), the matrix \( W \) is constructed by placing side by side the column vectors \( w_j \). Instead of computing the eigenvectors and eigenvalues of the covariance matrix \( WW^T \) with the size \( N \times N \), the the eigenvectors \( d_j \) and eigenvalues \( \mu_j \) of the covariance matrix \( WW^T \) with matrix of size \( M \times M \) are computed based on the following common theorem in linear algebra:

\[
W^T W d_j = \mu_j d_j
\]

By multiplying \( W \) to both sides of Eq. (11), we have:

\[
W^T W (W d_j) = \mu_j (W d_j)
\]

It means that the first \( M \) eigenvectors and eigenvalues of \( WW^T \) are given by \( W d_j \) and \( \mu_j \). The eigenvectors are sorted from high to low according to their corresponding eigenvalues. In fact, the smallest eigenvalue is associated with the eigenvector that finds the least variance. A facial image is projected on \( L (L << M) \) dimensions using the following equation:

\[
\Omega = \begin{bmatrix} v_1 & v_2 & \cdots & v_L \end{bmatrix}^T
\]

where \( v_i \) is the \( i^{th} \) coordinate of the facial image in the new space, which come to be the principle component.
3.2. MLPS Neural Network

The Multi-Layer Perceptrons (MLPs) have been assigned to implement the recognition core after feature extraction. As the improvement of original perceptrons, the MLPs have been upgraded by cascading one or more extra layers [10], called hidden layers, which are not directly connected to the external environment. The main algorithm of the MLPs learning is the propagation of the error backwards. This method can be shown briefly as the following steps:

- Correction of the output layer of weights using the following formula:

\[ w_{ho} = w_{ho} + (\eta \delta_h o_o) \]  

where \( w_{ho} \) is the weight connecting the hidden node \( h \) with the output node \( o \). The \( \eta \) denotes the learning rate and \( o_h \) is the output at the hidden node \( h \). The parameter \( \delta_h \) is expressed by the following equation:

\[ \delta_h = o_h (1 - o_h) (t - o_o) \]  

where \( o_h \) is the output at the node \( o \) of the output layer, and \( (t-o) \) describes the target output for that node.

- Correction of the input weights using the following equation:

\[ w_{ih} = w_{ih} + (\eta \delta_h o_i) \]  

where \( w_{ih} \) is the weight connecting the node \( i \) of the input layer with the node \( h \) of the hidden layer. The coefficient \( \eta \) expresses the learning rate, \( o_i \) is the input at the node \( i \) of the input layer. From Eq. (16), \( \delta_h \) is described as follows:

\[ \delta_h = o_h (1 - o_h) \sum_o (\delta_o w_{ho}) \]  

- For calculation of the error \( E \), the following function can be utilized by taking the average difference between the target and the output vector:

\[ E = \sqrt{\frac{1}{p} \sum_{n=1}^{p} (t_o - o_o)^2} \]  

where \( p \) is the number of nodes in the output layer, \( t_o \) and \( o_o \) are the target and output vectors at the output node \( o \).

In simulation, assume that 50 samples of 2 signs are used for each class, only 30 principle components corresponding to 30 eigenvectors having the largest eigenvalues are chosen for training process. It is noted that there are three networks generated to respond to three classes. In particular, the candidate sign will be identified an appropriate network only. The MLPs network (see Fig. 7), in this paper, has been designed so that the number of nodes in the output layer is equal to that of signs and the number of neurons in the input layer is corresponding to the number of eigenvectors. The training system is implemented based on the Mean Squared Normalized Error (MSE) performance function with some settings and parameters as listed in Table 3.

For each object, the node of an output layer will express two coefficients which are the sums of weights. By comparing the maximum of the two coefficients to the threshold, the current object is decided to be the two following conditions as the sign in the database:

- Maximum sum of weights is less than the threshold \( \rightarrow \) non-signs.
- Maximum sum of weights is greater than the threshold \( \rightarrow \) signs.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activation Function</td>
<td>Sigmoid</td>
</tr>
<tr>
<td>Performance Function</td>
<td>MSE</td>
</tr>
<tr>
<td>Number of hidden layers</td>
<td>2</td>
</tr>
<tr>
<td>Number of inputs</td>
<td>30</td>
</tr>
<tr>
<td>Number of outputs</td>
<td>2</td>
</tr>
<tr>
<td>Number of neurons in the first hidden layer</td>
<td>30</td>
</tr>
<tr>
<td>Number of neurons in the second hidden layer</td>
<td>2</td>
</tr>
<tr>
<td>Maximum Number of training Epochs</td>
<td>500</td>
</tr>
</tbody>
</table>

4. Experimental Results and Discussion

Simulation results of the proposed method for statistic images are represented. There are over 500 test images (150 for circle, 200 for square, and 150 for triangle) from urban areas and suburbs used for evaluation of 6 traffic signs with three classes. Fig. 8 shows the traffic signs which need to be recognized and Fig. 9 represents some original samples of recognition. In particular, the traffic signs were effectively recognized and shown on top left corner of each image.

In addition, more results are also shown in Table 4 with three events: success (recognized as the correct sign), missed (false in detection process), and false (recognized as the wrong sign). In practice, these images have been captured on Vietnam roads which have many different objects such as electric posts with wiring maze, street lights, or advertisement boxes can negatively affect image recognition. Moreover, the intermittent weather can be considered as a challenge in dangerous warning for drivers. In Table 4, it can be seen that the recognition method achieves the high performance based on the PCA-MLPs network, in which the effect of classification is very high. The fact is that the candidate sign is propagated in the 2-signs MLPs instead of the 6-signs MLPs network, that is, not only the accuracy rate is improved but also the time for identification is reduced significantly. Although some signs are missed in the segmentation process due to poor weather conditions, rules of the ratio of image size and the mean of correlations rejected unsuitable objects to save the resource for computation.
Figure 7. Structure of MLPs network for each class

Figure 8. Set of traffic signs used for evaluation

Figure 9. Result of simulation with some sample images
Table 4. Experimental results for images under complex conditions

<table>
<thead>
<tr>
<th>Class</th>
<th>Success</th>
<th>Missed</th>
<th>False</th>
<th>Recognition Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circle</td>
<td>144</td>
<td>5</td>
<td>1</td>
<td>96%</td>
</tr>
<tr>
<td>Square</td>
<td>182</td>
<td>13</td>
<td>5</td>
<td>91%</td>
</tr>
<tr>
<td>Triangle</td>
<td>143</td>
<td>4</td>
<td>3</td>
<td>95.3%</td>
</tr>
</tbody>
</table>

Based on the Scale-Invariant Feature Transform (SIFT), the Kus's method [15] only achieved the best performance when adding the color and orientation information. As a comparison, the proposed method is simpler than Kus's research when using less features for training. In order to achieve almost 90% of the accuracy with the SIFT feature extraction, Guo [16] represented the visual attention mechanism as a novel method. However, the drawback of Guo's method is that the performance is reduced in disturbance conditions. In this research, the PCA feature extraction of traffic sign images based on morphological classification for recognition using the MLPs is a novel method. The accurate rate of this method for Vietnam road sign is high and effective.

In the project of the Traffic Sign Recognition (TSR) system for images with poor condition images [17], the authors showed images with some clear signs about colours and shapes. Therefore, the simple feature extraction method of the images was applied using binary pixels and then they was recognized. This method can just have the high accuracy with the good image conditions. In this paper, we processed images with bad conditions and used from two chrome components of the YCbCr space to classify object types such circle, triangle and square. The PCA-MLPs method was employed to train and recognize a typical sign image on road with the high accuracy.

There have had a few publications related to the TSR using different feature extraction and recognition methods [18-20]. The authors applied different algorithms for traffic sign detection and recognition. All these algorithms are different from our proposed algorithm. In particular, from sign images with poor conditions of urban roads in Vietnam, the novel traffic sign recognition method is that detection and classification of sign images were morphologically applied based on calculation and correlations of the object images and then the PCA-MLPs was utilized for recognition of traffic signs.

5. Conclusions

In this research, a novel traffic sign recognition method was represented using the PCA-MLPs algorithm with two main issues of the morphological detection and classification. In particular, the sign objects were separately determined from two chrome components of the YCbCr space and then categorized into three classes: triangle, circle, and square by computing the correlations of these objects and samples in different rotated versions. The result is that the PCA-MLP-based recognition is implemented simpler and faster for the classification. In addition, the method was evaluated on statistic images captured on Vietnam roads with complex different conditions, but the high accuracy rate obtained is up to 96%. Simulation results have proved the effectiveness of the proposed approach.

REFERENCES


