Proceedings of the
2007 IEEE Intelligent Vehicles Symposium
Istanbul, Turkey, June 13-15, 2007

Invariant Road Sign Recognition with Fuzzy ARTMAP and Zernike Moments

Hasan Fleyeh, Mark Dougherty, Dinesh Aenugula, and Sruthi Baddam

Abstract—**In this paper, a novel approach to recognize road signs is developed. Images of road signs are collected by a digital camera mounted in a vehicle. They are segmented using colour information and all objects which represent signs are extracted, normalized to 36x36 pixels, and used to train a Fuzzy ARTMAP neural network by calculating Zernike moments for these objects as features. Sign borders and pictograms are investigated in this study.**

Zernike moments of sign borders and speed-limit signs of 210 and 150 images are calculated as features. A fuzzy ARTMAP is trained directly with features, or by using PCA for dimension reduction, or by using LDA algorithm as dimension reduction and data separation algorithm. Two Fuzzy ARTMAP Neural Networks are trained. The first NN determines the class of the sign from the shape of its border and the second one determines the sign itself from its pictogram. Training and testing of both NNs is done offline by using still images. In the online mode, the system loads the Fuzzy ARTMAP Neural Networks, and performs recognition process. An accuracy of about 99% is achieved in sign border recognition and 96% for Speed-Limit recognition.

I. INTRODUCTION

THE main goal of automatic sign recognition is to extract road signs from images of complex scenes under uncontrollable illumination. Road and traffic signs define a visual language that can be interpreted by drivers. They represent the current traffic situation on the road, show danger and difficulties facing drivers, give warnings to them, and help them with their navigation by providing useful information that makes the driving safe and convenient [1, 2].

Because of the complex environment of roads and the scenes around them, the detection and recognition of road and traffic signs may face some difficulties. The colour of the sign fades with time as a result of long exposure to sunlight, and the reaction of the paint with the air [1, 3]. Visibility is affected by weather conditions such as fog, rain, clouds and snow [1]. The colour information is very sensitive to the variations in the light conditions such as shadows, clouds, and the sun. [1, 3, 4]. It can be affected by the illuminant colour (daylight), illumination geometry, and viewing geometry [5]. Objects similar in colour to the road signs in the scene under consideration may be present, like buildings or vehicles. Signs may be found disoriented, damaged or occluded. If the image is acquired from a moving car, then it often suffers from motion blur and car vibration.

Research in road sign recognition is growing rapidly because of the real need for such systems in future vehicles. Aoyagi and Asakura [6] used neural networks for the recognition of the traffic sign. Vitabile et al. [7, 8] used a multi-layer perceptron neural network to classify road signs. de la Escalera et al. [9] used neural networks for classification of the traffic signs. It follows Adaptive Resonance Theory ART1. Fang et al. [1] carried out classification by the conceptual component module in which an ART2 network with a configurable long term memory to achieve classification.

This paper aims to present a new road sign recognition system based on using Zernike moments and Fuzzy ARTMAP neural networks as a shape and pictogram classifier. Two different Fuzzy ARTMAP neural networks are trained with Zernike moments computed for 210 and 150 images, respectively, to recognize and classify the signs. Speed-Limit signs are selected as case study for the classification stage because the detection and recognition of these signs could be combined by reading the vehicle's speed via a GPS and give a proper feedback to the driver to adjust vehicle speed. In addition, many such signs are present on the roads making data collection easy.

II. SWEDISH ROAD SIGNS

In contrast to many other European countries, Swedish road signs are characterized by using yellow background colour for Warning, and Prohibitory signs. Swedish traffic signs can be categorized into four types:

- Warning signs: they are red rimmed yellow triangle with symbol or letter messages.
- Prohibitory signs: they are red rimmed yellow or blue circle with different symbols or messages. An octagon is used for stop sign.
- Regulatory signs or mandatory signs: they are circle shaped with blue filling colour and white symbols or arrows.
- Indicatory and supplementary signs: They are characterised by using rectangles with different background colours such as yellow, green, or blue etc. with white or black symbols or messages.

Figure 1 shows the basic grouping of Swedish road signs depending on the colour of the border. They can be divided,

Hasan Fleyeh, Computer Science Department, Dalarna University, 78188 Borlänge, Sweden; e-mail: hfl@du.se. Mark Dougherty, Computer Science Department, Dalarna University, 78188 Borlänge, Sweden; e-mail: mdo@du.se. Dinesh Aenugula, Computer Science Department, Dalarna University, 78188 Borlänge, Sweden; e-mail: v04draen@du.se, and Sruthi Baddam Computer Science Department, Dalarna University, 78188 Borlänge, Sweden; e-mail: v04sruba@du.se.

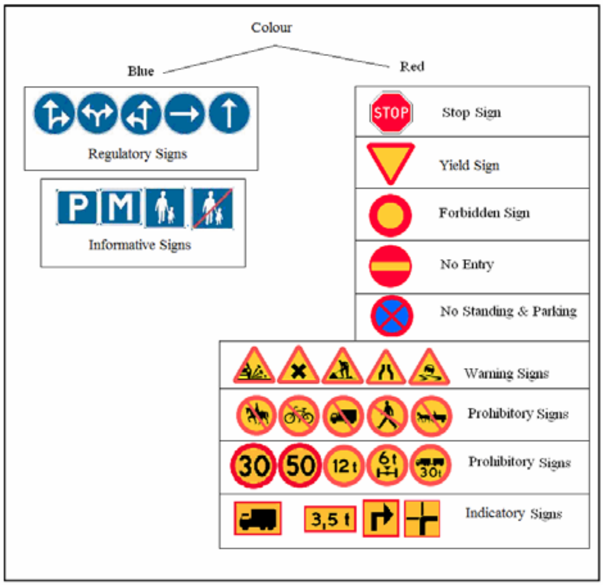

Fig. 1. Main colour –shape combinations of Swedish road signs.

mainly, into two major groups, *red* and *blue*. This paper concentrates on the recognition of the signs in the red border group followed by the recognition of Speed-limit signs. Speed-Limit signs are part of the prohibitory signs. There are five standard Speed-limit signs; namely 30, 50, 70, 90, and 110 km/h. However, there are many other special Speed-Limit signs such as 5, 10, 15, 20, 40 km/h. These special Speed-Limit signs are occasionally found in use and it is therefore hard to collect sufficient examples for training and testing the system.

III. ZERNIKE MOMENTS

Zernike moments of order p with repetition q of a discrete binary image with image intensity function $f(x,y)$ inside a unit circle is given by[10]

$$Z_{pq} = \frac{p+1}{\pi} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} R_{pq}(r_{xy}) e^{-jq\theta_{xy}} f(x,y) \tag{1}$$

where $0 \le r \le 1$, $r_{xy} = \sqrt{x^2+y^2}$, and $\theta_{xy} = \tan^{-1}(y/x)$.

$R_{pq}(r_{xy})e^{-jq\theta_{xy}}$ are the Zernike polynomials and $R_{pq}(r)$ is the radial polynomial. $R_{pq}(r)$ is defined as [9]

$$R_{pq}(r) = \sum_{k=0}^{(p-|q|/2)} (-1)^k \frac{(p-k)!}{k!((p+|q|)/2-k)!((p-|q|)/2-k)!} r^{p-2k} \tag{2}$$

where $0 \le |q| \le p$ and $p-|q|$ is even.

In any recognition problem, the images might not have the same size, position or orientation. But the features extracted for similar images must be the same irrespective of these parameters, which means that the features extracted from the images are invariant under translation, rotation and scaling.

Zernike moments are rotational invariant by nature which means that the magnitude of the Zernike moments of an image remains the same even if the phase angle of that image is changed.

Moment-based features of an object are said to be translation invariant if they remain the same even when the object is located in different positions in the image. Translation invariance of Zernike moments are achieved by moving the origin of the image to the centroid of the object. This is done by transforming the image $f(x,y)$ into another image $f(x+x_{cen}, y+y_{cen})$ where (x_{cen}, y_{cen}) are object centroids.

Scale invariance is obtained by making the area of the object under consideration the same for all images and hence, moments of an image are said to be scale invariant if their values do not change when image size is changed.

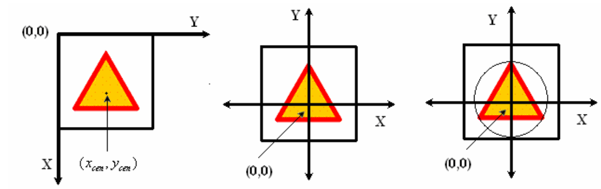

Fig. 2. Main steps to perform translations and scaling invariance.

To calculate modified Zernike moments which are rotation, translation and scale invariant, the following steps are performed:

1. To carry out the translation invariance, the binary image is normalised by calculating object's area and centroid (x_{cen}, y_{cen}) from the following equations:

$$a = \sum_x \sum_y f(x,y) \tag{3}$$

$$x_{cen} = \frac{1}{a} \sum_x x f(x,y) \tag{4}$$

$$y_{cen} = \frac{1}{a} \sum_y y f(x,y) \tag{5}$$

2. Calculate objects' pixel coordinates of the new image shown in figure 2:

$$x' = y - y_{cen} \quad , \quad y' = x_{cen} - x \tag{6}$$

where (x', y') are the transformed coordinates and (x,y) are the original coordinates of the object.

3. Find the radius of the minimum circle containing the object by calculating the furthest object's pixel from the centroid (x_{cen}, y_{cen}), denoted $r_{\max}$, using Euclidean distance.

$$r_{\max} = \sqrt{(x' - x_{cen})^2 + (y' - y_{cen})^2} \tag{7}$$

Use $r_{\max}$ to map object's coordinates to the modified coordinates.

4. Map the coordinates of every object's pixel to be within a unit circle by calculating

$$x'' = x' / r_{\max} \quad , \quad y'' = y' / r_{\max} \tag{8}$$

5. Calculate Zernike moments using x'' and y'' values achieved from the former step.

6. Calculate the modified Zernike moments z'_{pq} by dividing the Zernike moments by the area of the object.

$$z'_{pq} = z_{pq} / a \qquad (9)$$

IV. SYSTEM OVERVIEW

The proposed system, shown in figure 3, consists of certain number of units which work together to perform the recognition of road signs. These units are as follows:

a. The Camera:
Data acquisition of images for training, testing and for real-time applications is carried out by a camera mounted on a moving vehicle. More than 3000 images collected in different light conditions are used for training and testing of the algorithms.

b. Colour Segmentation:
Colour segmentation is an important step to eliminate all background objects and unimportant information in the image. It generates a binary image containing the road signs and any other objects similar to the road sign in colour. Colour segmentation is carried out by a shadow and highlight invariant algorithm [11].

c. Shape Analysis:
The output of the former unit is a binary image with a number of objects that could be probable road signs. This unit is used in two modes, the offline and the online mode. In the offline mode is it invoked to create or update the image data base. Objects in the segmented image are extracted using the connected components labelling algorithm, size normalized to 36x36 pixels and saved in the database for the training of the classifier. The following set of equations is used for normalization:

$$x' = N(x - x_{\min}) / (x_{\max} - x_{\min}) \qquad (10)$$

$$y' = N(y - y_{\min}) / (y_{\max} - y_{\min}) \qquad (11)$$

Where the coordinates values $x_{\min}$, $x_{\max}$, $y_{\min}$, $y_{\max}$ are the rectangle vertices containing the sign before normalization with sides parallel to the vertical and horizontal axes, and (x', y') are the coordinates of a generic point in the new NxN matrix corresponding to the (x, y) coordinates of the pixel of the original matrix. In the online mode, the same aforementioned process is followed, but images are forwarded to the feature extraction unit instead.

d. Training Database
The training data base consists of 460 binary images of size 36x36 pixels used for training of the Fuzzy ARTMAP by calculating Zernike moments. The database comprises 210 images of border shapes and 150 images for Speed-Limit signs. It is created by the method described in the preceding step. Figures 4 and 5 show part of the database of the borders and pictograms used for recognition of the speed-limit signs.

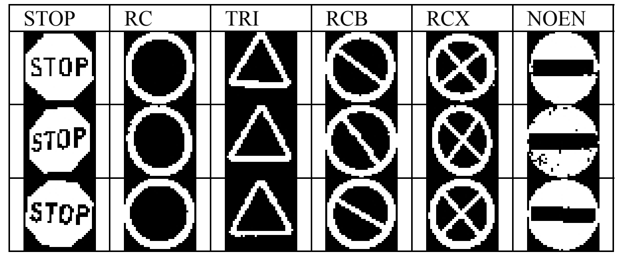

Fig. 4. Part of the training set for the Border recognition.

e. Feature Extraction:
This unit contains feature extraction algorithm such as Zernike moments calculation algorithm and some other algorithms used for dimension reduction such as Principal Components Analysis (PCA) and Linear Discriminant Analysis (LDA).

f. Fuzzy ARTMAP Classifier:
Fuzzy ARTMAP belongs to the family of neural networks which originally developed from the competitive learning paradigm to overcome the stability- plasticity dilemma. This family is called Adaptive Resonance Theory or ART. Fuzzy ARTMAP is capable of performing fast and stable supervised learning both in analogue and binary patterns. Comparisons carried out by Heinke and Hamker [12] showed that Fuzzy ARTMAP is superior to the other neural networks classifiers. It requires less time to converge to a solution due to the small number of training epochs required by the network to learn the input data. Furthermore, Fuzzy ARTMAP is even faster than other ARTMAP techniques due to the computationally cheap Input/Output mapping. The other strong point of Fuzzy ARTMAP is that its classification result can be easily interpreted; IF-THEN rules can be readily derived from a Fuzzy ARTMAP.
The architecture of Fuzzy ARTMAP is described in details by Carpenter et al. [13]

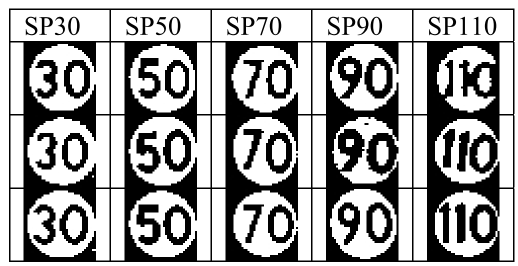

Fig. 5. Part of the training set for the Interior of the sign.

V. EXPERIMENTS AND RESULTS

As mentioned in section 2, the Swedish road signs use yellow as background colour for Warning and Prohibitory signs, therefore, speed-limit sign detection (figure 6) is achieved by searching the image for a red object containing yellow in its convex-hull. Once such object is found, a check to establish whether it is probably a sign is carried out. This operation is achieved in two stages. In the first stage, the red border is recognized, followed by the recognition of the interior part of the sign or pictogram, figure 6.

Fig. 6. Dissection of Swedish speed limit sign.

Shapes of red rimmed road signs can be divided into seven categories. Because Zernike moments are used as features, and since these moments are rotation invariant by definition, it is impossible to discriminate “upward” and “downward” triangles. Therefore, number of classes is reduced to six by merging the two triangle classes, figure 4. The classification process is carried out using three methods:

Method 1:

For shape recognition stage, features extraction is achieved by using Zernike moments and classification using Fuzzy ARTMAP. Zernike moments of the normalized images are extracted. The 40 feature dataset is directly passed to the fuzzy ARTMAP. The fuzzy ARTMAP is trained and tested by two different sets of images which have been selected randomly. In the training phase, a classification rate of 98% is achieved compared with 96.6% for testing; this is shown in table 1.

For speed-limit recognition stage, the same aforementioned procedure is followed. As shown in table 2, a classification rate of 93.7% and 76% is achieved for training and test sets, respectively.

Method 2:

Zernike moments of sign borders are calculated as features, PCA is invoked as feature reduction algorithm and classification is carried out using Fuzzy ARTMAP. The initial number of Zernike moments of the images is 40. It is then reduced by PCA. Fuzzy ARTMAP is trained and tested using two different sets of randomly selected images. According to table 1, a classification rate of 94.7% is achieved for training set, while 90% is achieved for test set.

A similar procedure is followed for speed-limit signs. As shown in table 2, for the training set a classification rate of 88.3% is achieved, compared with 48% for test set.

Method 3:

In this method, Zernike moments are computed as features, LDA is used for dimension reduction and data separation is achieved by increasing the inter-class distance, Fuzzy ARTMAP is used as classifier. For shape recognition, a classification rate of 99% is achieved for training set compared with 100% for the test set, table 1. For speed-limit signs, as shown in table 2, a classification rate of 98.5% is achieved for training set and 96% for test set.

Figures 7 and 8 show the scatter plots of the above mentioned methods for both training and test sets. It is very clear from these figures that method 3 is the best among the methods because of the increase in the inter-class distance achieved by the LDA algorithm.

VI. CONCLUSIONS

In this paper, a new method to classify road signs is presented. The method depends mainly on using Zernike moments as invariant features. Zernike moments are invariant to rotation by definition. Furthermore, a method to make Zernike moments invariant to scaling and translation is shown in this paper. Invariance is an important property to deal with images of different transformations in the image plane, which is very likely to happen when dealing with traffic signs. Two types of dimension reduction techniques are tested; PCA and LDA. It is shown that LDA gives better results compared with PCA. It is also noticed that the general performance of the suggested algorithm is better with shape recognition compared to speed-limit; this is an expected result because the speed-limit signs are more similar to each other compared with sign borders.

For future work, more features or feature fusion will be tested. Legendre moments and Orthogonal Fourier-Mellin descriptors are planned to be tested in the future too.

REFERENCES

[1] C. Fang, C. Fuh, S. Chen, and P. Yen, "A road sign recognition system based on dynamic visual model," presented at The 2003 IEEE Computer Society Conf. Computer Vision and Pattern Recognition, Madison, Wisconsin, 2003.

[2] C. Fang, S. Chen, and C. Fuh, "Road-sign detection and tracking," *IEEE Trans. on Vehicular Technology*, vol. 52, pp. 1329-1341, 2003.

[3] J. Miura, T. Kanda, and Y. Shirai, "An active vision system for real-time traffic sign recognition," presented at 2000 IEEE Intelligent Transportation Systems, Dearborn, MI, USA, 2000.

[4] S. Vitabile, G. Pollaccia, G. Pilato, and F. Sorbello, "Road sign Recognition using a dynamic pixel aggregation technique in the HSV color space," presented at 11th Inter. Conf. Image Analysis and Processing, Palermo, Italy, 2001.

[5] S. Buluswar and B. Draper, "Color recognition in outdoor images," presented at Inter. Conf. Computer vision, Bombay, India, 1998.

[6] Y. Aoyagi and T. Asakura, "A study on traffic sign recognition in scene image using genetic algorithms and neural networks," presented at The1996 IEEE IECON 22nd Inter. Conf. on Industrial Electronics, Control and Instrumentation, Taipei, Taiwan, 1996.

[7] S. Vitabile, A. Gentile, G. Dammone, and F. Sorbello, "Multi-layer perceptron mapping on a SIMD architecture," presented at The 2002 IEEE Signal Processing Society Workshop, 2002.

[8] S. Vitabile, A. Gentile, and F. Sorbello, "A neural network based automatic road sign recognizer," presented at The 2002 Inter. Joint Conf. on Neural Networks, Honolulu, HI, USA, 2002.

[9] A. de la Escalera, J. Armingol, and M. Mata, "Traffic sign recognition and analysis for intelligent vehicles," *Image and Vision Comput.*, vol. 21, pp. 247-258, 2003.

[10] N. Kamila, S. mahapatra, and S. Nanda, "Invariance image analysis using modified Zernike moments," *Pattern Recognition Letters*, vol. 26, pp. 747-753, 2005.

[11] H. Fleyeh, "Shadow And Highlight Invariant Colour Segmentation Algorithm For Traffic Signs," presented at 2006 IEEE Conf. on Cybernetics and Intelligent Systems, Bangkok, Thailand, 2006.

[12] D. Heinke and F. Hamker, "Comparing neural networks: A benchmark on Growing Neural Gas, Growing Cell Structures and Fuzzy ARTMAP," *IEEE Transactions on Neural Networks*, vol. 9, pp. 1279-1291, 1998.

[13] G. Carpenter, S. Grossberg, N. Markuzon, J. Reynolds, and D. Rosen, "Fuzzy ARTMAP: a neural network architecture for Incremental supervised learning of analog multidimensional maps," *IEEE Trans. Neural Networks*, vol. 3, pp. 698-713, 1992.

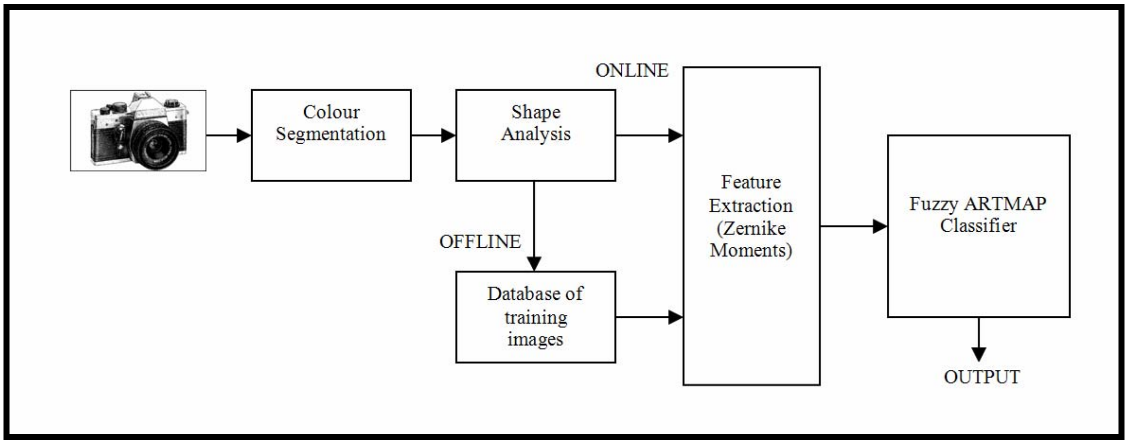

Fig. 3. Block Diagram of the Proposed System.

TABLE I
TRAINING AND TEST SETS ERROR REPORT FOR THE THREE DIFFERENT METHODS USED FOR SHAPE CLASSIFICATION

Method / Label	Classification with No Reduction		Classification with PCA		Classification with LDA	
	Training%	Test %	Training%	Test %	Training%	Test %
STOP	97.6	100	98.4	100	99.2	100
RC	97.6	100	95.2	60.0	98.4	100
TRI	99.2	100	96.8	100	99.2	100
RCB	98.4	100	88.8	100	99.2	100
RCX	96.8	80.0	93.4	80.0	99.2	100
NO ENTRY	98.4	100	96.0	100	99.2	100
Avg.	**98.0**	**96.6**	**94.7**	**90.0**	**99.0**	**100**

TABLE II
TRAINING AND TEST SETS ERROR REPORT FOR THE THREE DIFFERENT METHODS USED FOR SPEED- LIMIT CLASSIFICATION

Method / Label	Classification with No Reduction		Classification with PCA		Classification with LDA	
	Training%	Test%	Training%	Test%	Training%	Test t%
SP30	94.4	80.0	82.4	40.0	97.6	100
SP50	87.2	40.0	85.6	60.0	98.4	100
SP70	94.4	80.0	87.2	0.00	99.2	100
SP90	94.4	80.0	95.2	80.0	99.2	100
SP110	98.4	100	91.2	60.0	98.4	80.0
Avg.	**93.7**	**76.0**	**88.3**	**48.0**	**98.5**	**96.0**

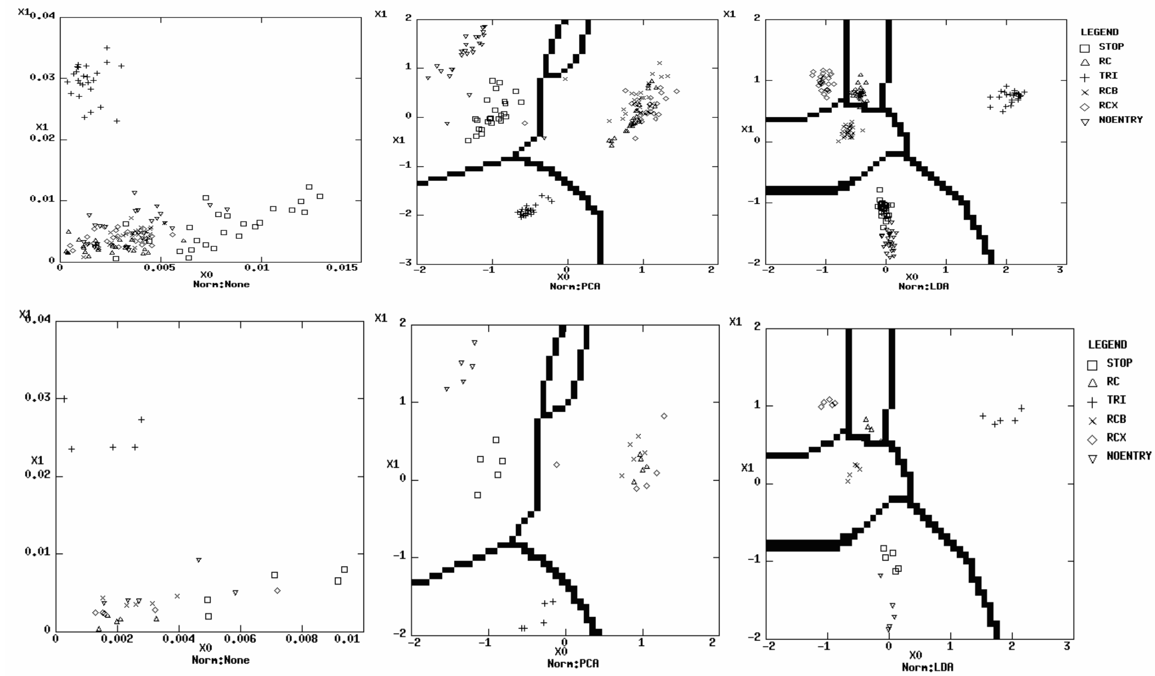

Fig. 7. Scatter error graphs of the training and test sets for shape recognition.

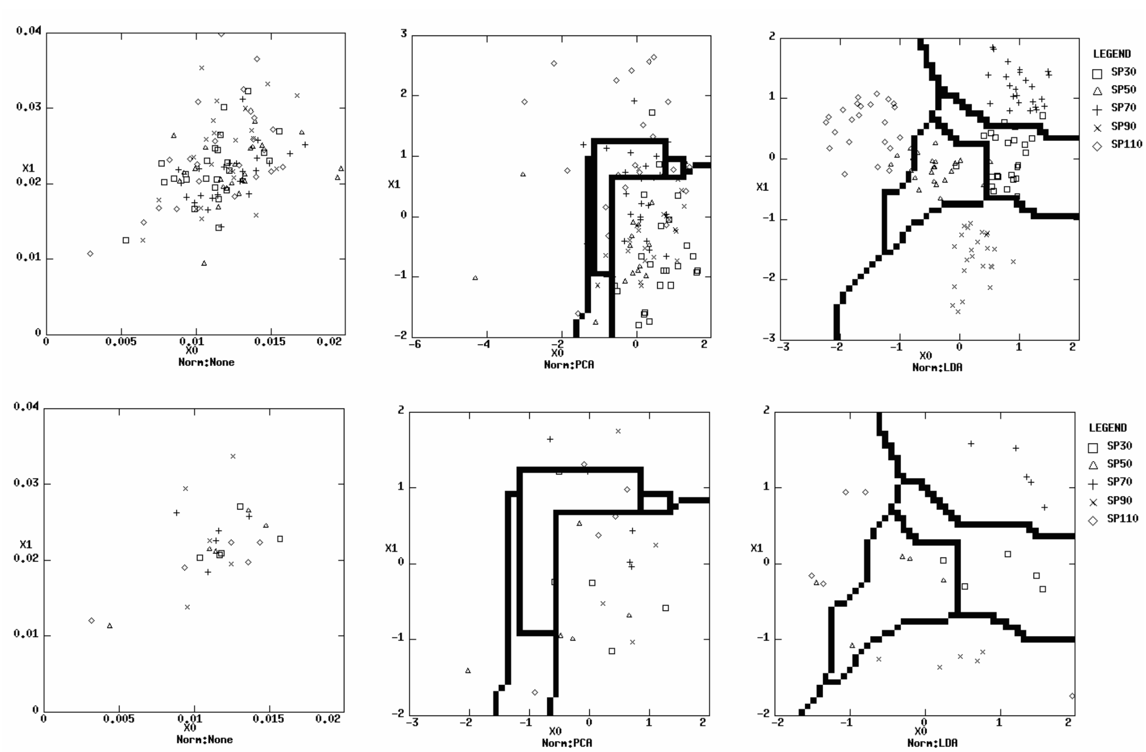

Fig. 8. Scatter error graphs of the training and test sets for speed-limit recognition.