International Journal of Soft Computing 3 (4): 302-307, 2008
ISSN: 1816-9503
© Medwell Journals, 2008

Neural Classifier System for Object Classification with Cluttered Background Using Invariant Moment Features

[1]B. Nagarajan and [2]P. Balasubramanie
[1]Department of Computer Applications, Bannari Amman Institute of Technology, Sathyamangalam, Erode District, Tamil Nadu, India
[2]Department of Computer Science and Engineering, Kongu Engineering College, Perundurai, Erode District, Tamil Nadu, India

Abstract: Object recognition and classification in a multi-environment is an important part of machine vision. The goal of this study, is to build a system that classifies the objects amidst background clutter and mild occlusion. This study addresses the issues to classify objects of real-world images containing side views of cars with cluttered background with that of non-car images with natural scenes. The threshold technique with background subtraction is used to segment the background region to extract the object of interest. The background segmented image with region of interest is divided into rectangular sub-images of equal size. The moment features which are invariant to Rotation, Scaling and Translation (RST) are extracted from each rectangular block. The features of the objects are fed to the back-propagation neural classifier. Thus, the performance of the neural classifier is compared with various categories of rectangular block size. Quantitative evaluation shows improved results of 84.9%. A critical evaluation of our approach under the proposed standards is presented.

Key words: Object classification, background segmentation, invariant moment, neural classifier

INTRODUCTION

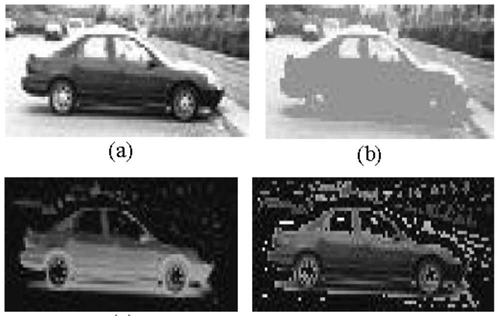
In computer vision, object classification plays a major role in applications such as security systems, traffic surveillance system, target identification, etc. The classification system faces 2 types of problem. Objects of same category with large variation in appearance. The objects with different viewing conditions like occlusion, complex background containing buildings, clouds, trees, road views, etc. The objects of interest to be classified are considered to be present in still, gray-scale images. This study, tries to bring out the importance of the background elimination with moment based feature extraction method of varying sub-block size for object classification. Thus images are divided into a rectangular sub-blocks instead of standard squared sub-blocks. The objects of interest being a car and non-car images are classified in this study containing real world data set.

Image understanding is a major area where researchers design computational systems that can identify and classify objects automatically. Identification and classification of vehicles has been a focus of investigation over last decades (Hsieh *et al*., 2006; Shan *et al*., 2005; Sun *et al*., 2006). A new approach to object detection that makes use of a sparse, part-based representation is proposed by Agarwal *et al*. (2004). This study gives very promising results in the detection of vehicles from a group of non-vehicle category of natural scenes. Moment invariants are important shape descriptors in machine vision especially in the field of object detection and classification. One of the most popular and widely used contour-based shape descriptors is a set derived by Hu Paschalakis and Lee (1999) presents the issue of classification among objects which have identical shapes, using grey level images based on invariant moments. Yinan *et al*. (2003) proposes new formulas of moment invariants that are defined united moments.

BACKGROUND REMOVAL AND MAPPING FUNCTION

The overall complexity increases for the natural images as the object of interest is lying on the background region. In object classification problem, it is essential to distinguish the object of interest and the

Corresponding Author: B. Nagarajan, Department of Computer Applications, Bannari Amman Institute of Technology, Erode District, Sathyamangalam-638 401, Tamil Nadu, India



(a) (b) (c) (d)

Fig. 1: a) A sample image with natural background denoted as I(x,y). b) The small regions are removed by filling the holes. c) Image difference obtained by subtraction (a) by (b) denoted as d(x,y). d) Image obtained by mapping function f(x,y)

background. Segmentation of object is done through background subtraction technique. This method is more suitable when the intensity levels of the objects fall outside the range of levels in the background.

An object with natural background is shown in Fig. 1. Initially morphological operations are applied to suppress the residual errors with help of open and close pair statements (Li *et al*., 2004; Richord *et al*., 2005). The small regions are removed by filling the holes. Then the image subtraction is applied with the previous result. Thus the object is segmented from the background.

A mapping function (1) is used to restore the object of interest from that of the subtracted image.

$$f(x,y) = \begin{cases} 0, \text{ if } d(x,y)=0 \\ I(x,y), \text{ Otherwise} \end{cases} \quad (1)$$

where, f(x,y) is the transformed image, d(x,y) is image difference after fill operation and I(x,y) is the original image.

INVARIANT MOMENTS FEATURES

Moment-based invariants are the most common region-based image invariants which have been used as pattern features in many applications (Rizon *et al*., 2006; Xu and Li, 2008; Devendran *et al*., 2008). Hu first introduced a set of invariants using non-linear combinations based on regular moments.

Regular moments are defined as:

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)\,dx\,dy, \; p,q = 0,1,2,.... \quad (2)$$

where, m_{pq} is the (p + q) th order moment of the continuous image function f(x, y).

The central moments of f(x, y) are defined as:

$$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\bar{x})^p (y-\bar{y})^q f(x,y)\,dx\,dy \quad (3)$$

Where,

$$\bar{x} = m_{10}/m_{00}$$

and

$$\bar{y} = m_{01}/m_{00}$$

which are the centroid of the images.

The central moments are computed using the centroid of the image, which is equivalent to the regular moments of an image whose center has been shifted to coincide with its centroid; therefore the central moments are invariant to image translations.

The matrix for image scale change is:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (4)$$

To obtain scale invariance, let $f'(x', y')$ represent the image f(x, y) after scaling the image by $S_x = S_y = \alpha$, so $f'(x', y') = f'(\alpha x, \alpha y') = f(x, y)$ and $x' = \alpha x$, $y = \alpha y$, then we can easily prove:

$$m'_{pq} = \alpha^{p+q+2} m_{pq} \quad (5)$$

Similarly,

$$\mu'_{pq} = \alpha^{p+q+2}\mu_{pq,}\, \mu'_{00} = \alpha^2 \mu_{00} \quad (6)$$

We can define normalized central moments as:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}, \; \gamma = (p+q+2)/2, \; p+q = 2,3,... \quad (7)$$

μ_{pq} is invariant to changes of scale because:

$$\eta'_{pq} = \frac{\mu'_{pq}}{\mu'^{\gamma}_{00}} = \frac{\alpha^{p+q+2}\mu_{pq}}{\alpha^{2\gamma}\mu^{\gamma}_{00}} = \frac{\mu_{pq}}{\mu^{\lambda}_{00}} = \eta_{pq} \quad (8)$$

Based on normalized central moments, the following orthogonal (and translational) invariants have been derived by Hu.

$$
\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
&\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + (3\eta_{21} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned} \tag{9}
$$

The central moments are independent of the position and image size. The absolute orthogonal moment invariant is orientation independent. We combine the absolute orthogonal moment invariant with the similitude invariants of central moments. Then classification can be independent of position, size and orientation.
Features extraction consists of the following:

- Vector of length seven containing central moments made invariant to similitude transforms.
- Vector of length seven containing central moments made orthogonally invariant.
- Vector of length seven containing invariants with respect to both similitude and orthogonal transformations.

Combining all the three vectors, the feature space is populated with 21 features containing seven in each category.

BUILDING A NEURAL CLASSIFIER

A binary Artificial Neural Network (ANN) classifier is built with back-propagation algorithm that learns to classify an image as a member or nonmember of a class.

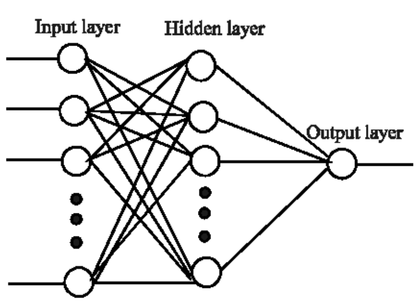


Fig. 2: The Three layer neural architecture

The number of input layer nodes is equal to the dimension of the feature space obtained from the moments and invariant features. The number of output nodes is usually determined by the application (Khotanzand and Chung, 1998) which is 1 (either "Yes/No") where, a threshold value nearer to 1 represents "Yes" and a value nearer to 0 represents "No". The neural classifier is trained with different choices for the number of hidden layer. The final architecture is chosen with single hidden layer shown in Fig. 2 that results with better performance.

The connections carry the outputs of a layer to the input of the next layer have a weight associated with them. The node outputs are multiplied by these weights before reaching the inputs of the next layer. The output neuron (10) will be representing the existence of a particular class of object.

$$
O_j^l(k) = f\left(\sum_{m=0}^{N l-1} w_{jm}^l O_m^{l-1} \right) \tag{10}
$$

PROPOSED WORK

This study addresses the issues to classify objects of real-world images containing side views of cars amidst background clutter and mild occlusion. The objects of interest to be classified are car (positive) and non-car (negative) images taken from University of Illinois at Urbana-Champaign (UIUC) standard database. The image data set consists of 1000 real images for training and testing having 500 in each class. The sizes of the images are uniform with the dimension 100×40 pixels.

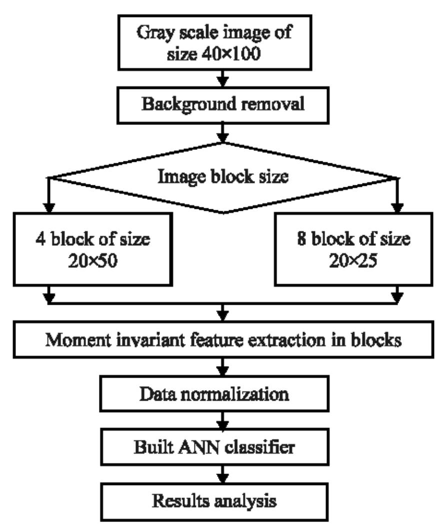


Fig. 3: The description of the proposed work

The proposed framework consists of two methods followed by background removal as given in the study Method-I: 4 Blocks of size 20×50, Method-II: 8 Blocks of size 20×25. Twenty one invariant moment features are calculated from each single block of sub-image using equations mentioned in the study. Zscore normalization is applied for the moment features which are the deviation from its mean by standard deviation. This process improves the performance of the neural classifier. The overall flow of the framework is shown in Fig. 3.

IMPLEMENTATION

We trained our methods with different kinds of cars against a variety of background, partially occluded cars of positive class. The negative training examples include images of natural scenes, buildings and road views. The training is done with 500 images (250 positive and 250 negative) against both the methods. The testing of images are done with 1000 images (500 positive and 500 negative) taken from the same image database.

The feed-forward network for learning is done for both method-I and method-II. The input nodes for method-I is 84 (4 blocks × 21 features) and for method-II is 168 (8 blocks × 21 features). Optimal structure validation is done and the structure given in Fig. 2, performs well and leads to better results. Thus the optimal structure of the neural classifier for method-I is 84-5-1 and for method-II is 168-20-1.

Table 1: Parameters for training of the neural classifier

Parameters	Method-I	Method-II
Learning rate	0.5	0.5
Performance goal	0.01	0.01
No. of epochs taken to meet the performance goal.	787	3041
Time taken to learn	9.156 Secs	53.31 Secs

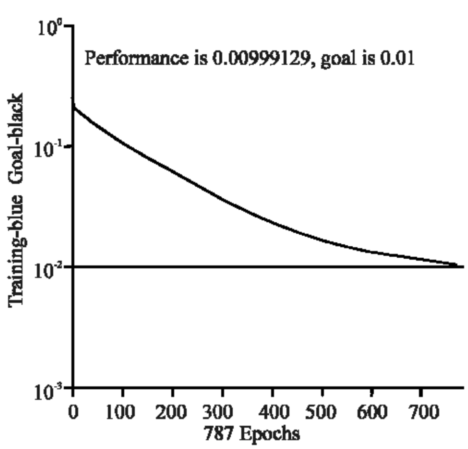


Fig. 4: The performance graph of neural network training for Method-I: 4 Blocks of size 20×50

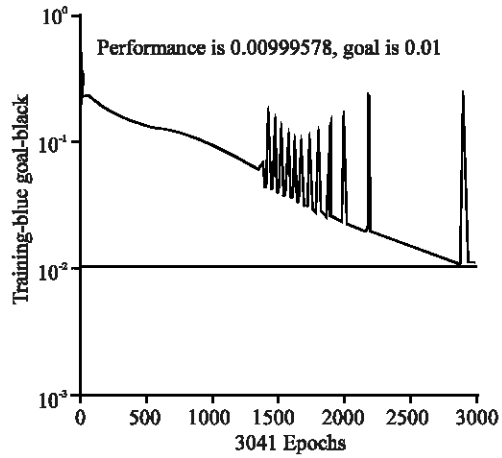


Fig. 5: The performance graph of neural network training for Method-II: 8 Blocks of size 20×25

The various parameters for the neural classifier training for both the methods are given in Table 1. The performance graph of the neural classifier for method-1 and method-II are shown in Fig. 4 and 5, respectively.

RESULTS AND DISCUSSION

In object classification problem, the four quantities of results category are given:

- True Positive (TP):
  Classify a car image into class of cars.
- True Negative (TN):
  Misclassify a car image into class of Non-cars.
- False Positive (FP):
  Classify a non-car image into class of non-cars.
- False Negative (FN):
  Misclassify a non-car image into class of cars.

The objective of any classification is to maximize the number of correct classification denoted by True Positive Rate (TPR) and False Positive Rate (FPR) where by minimizing the wrong classification denoted by True Negative Rate (TNR) and False Negative Rate (FNR).

$$\text{TPR} = \frac{\text{Number of true positive (TP)}}{\text{Total number of positive in data set (nP)}}$$

$$\text{TNR} = \frac{\text{Number of true negative (TN)}}{\text{Total number of negative in data set (nN)}}$$

$$\text{FPR} = \frac{\text{Number of false positive (FP)}}{\text{Total number of positive in data set (nP)}}$$

Table 2: Comparison of experimental methods

Threshold for classification: 0.7	Classifying positive images (Car images)		Classifying negative images (Non-car images)	
	TPR	TNR	FPR	FNR
Method-I				
4 Blocks of size 20×50	74.0%	26.0%	80.8%	19.2%
	Method-I Overall classification accuracy (TPR+FPR)/2 is 77.4%			
Method-II				
8 Blocks of size 20×25	79.6%	20.4%	90.2%	9.8%
	Method-II Overall Classification Accuracy (TPR+FPR)/2 is 84.9%			

Fig. 6: Sample results of the neural classifier of the category car images with cluttered background and mild occlusion

Fig. 7: Sample results of the neural classifier of the category non-car images containing trees, road view, bike, wall, buildings and persons

$$\mathrm{FNR} = \frac{\text{Number of false negative (FN)}}{\text{Total number of negative in data set (nF)}}$$

The values of nP and nN used as testing samples are 500 and 500, respectively. Most classification algorithm includes a threshold parameter for classification accuracy which can be varied to lie at different trade-off points between correct and false classification. The comparison of results of the proposed methods is shown in Table 2 which is obtained with an activation threshold value of 0.7.

Classified images of category car and non-car as resultant sample images are shown below in the Fig. 6 and 7, respectively.

It is evident that the classifier with 8 blocks of size 20×25 (Method-II) is showing improved overall results of 84.9% of classification accuracy compared with that of 4 blocks of size 20×50 (Method-I).

CONCLUSION

Thus an attempt is made to build a system that classifies the objects amidst background clutter and mild occlusion is achieved to certain extent. Thus the goal to classify objects of real-world images containing side views of cars with cluttered background with that of non-car images with natural scenes is presented. Comparing the results in Table 2, the performance of the proposed method with 8 blocks of size 20×25 (rectangular blocks) with invariant moment features after background removal gives a satisfactory classification rate of 84.9%. The limitation of this method is the object with a high degree of occlusion for classification. Further work extension can be made to improve the performance of the classifier system with the inclusion of feature selection process. This complete work is implemented using neural network and image processing toolbox in Matlab 6.5.

REFERENCES

Agarwal, S., A. Awan and D. Roth, 2004. Learning to detect objects in images via a sparse, part-based representation. IEEE. Trans. Pattern Anal. Machine Intell., 26 (11): 1475-1490.

Devendran, V. *et al*., 2008. Invariant Moments to Scene Categorization Using Support Vector Machines, 2: 128-133.

Hsieh, J.W. *et al*., 2006. Automatic Traffic surveillance system for vehicle tracking and classification. IEEE. Trans. Intell. Trans. Sys., 7 (2): 175-187.

Khotanzand, A. and C. Chung, 1998. Application of Multi-Layer Perceptron Neural Networks to Vision Problem. Neural Computing and Applications, Springer-Verlag London Limited, pp: 249-259.

Li, L. *et al*., 2004. Statistical modeling of complex backgrounds for foreground object detection. IEEE. Trans. Image Proc., 13 (11): 1459-1472.

Paschalakis, S. and P. Lee, 1999. Pattern recognition in grey level images using moment based invariant features. IEEE. Conf. Image Proc. Applied, 465: 245-249.

Richord, J.R. *et al*., 2005. Image change detection algorithms: A systematic survey. IEEE. Trans. Image Proc., 14 (3): 294-306.

Rizon, M. *et al*., 2006. Object Detection using Geometric Invariant Moment. Am. J. Applied Sci., 6: 1876-1878.

Shan, Y. *et al*., 2005. Vehicle Identification between Non-Overlapping Cameras without Direct Feature Matching. 10th IEEE. Int. Conf. Comput. Vision, pp: 1550-1558.

Sun, Z., G. Bebis and R. Miller, 2006. Monocular precrash vehicle detection: Features and classifiers. IEEE. Trans. Image Proc., 15: 2019-2034.

Xu, D. and H. Li, 2008. Geometric moment invariants. Pattern Recog., 41: 240-249.

Yinan, S. *et al*., 2003. United Moment Invariants for Shape Discrimination. IEEE Int. Conf. Robotics, Intell. Sys. Sig. Proc., pp: 88-93.