Moment based Object Recognition using Artificial Neural Network

Saibal Dutta and P.K.Nanda
Image processing and computer vision Lab,
Department of Electrical Engineering,
National Institute of Technology, Rourkela-769008, Orissa, India
saibal.13@gmail.com, pknanda_d13@rediffmail.com

***Abstract*— In this paper, we propose a Artificial Neural Network(ANN) based scheme for Object recognition.The invariance properties of geometric moments as well as lower order moments corresponding to partially occluded objects are used to train a feedforward ANN. The trained neural network is used to predict the actual moments from the moments of the object with different occlusions. The object is recognized based on the comparison of the actual moments and the predicted moments.**

I. Introduction

Moments and their invariance properties have been successfully used in different pattern recognition tasks [1-4]. There has been substantial research effort to use the invariant moments for recognition of objects in different tasks of image analysis. The moment invariance to scalling, rotation and translation properties is first proposed by Hu[1], who has demostrated the application of invarints of geometric moments in image analysis. The modified version of the invarince theorems have been presented by Reiss[2] and corroborted the use of invarince properties for recognition. The invariance properties of the moments do not hold good for unequal scalling of images and hence a method has been proposed [3] to preserve the invariance under above condition. In order to deal with blurred images Z.Tianxu et al [4] has proposed moment invariants approach that is not affected by blur, rotation, scale and translation of images. The image representation ability, noise sensitivity, and information redundancy of different moments have been investigated by C.Teh et al [5] and the relationships among different moments are also establised. It has been reported in the litrature that Zernike and Pseudo-Zernike moments are very much useful while reconstructing the images and rotation invariance feature has been used for object recognition[6]. Recently, a set of orthogonal, noise robust transformation invariants, distribution sensitive moments called as Eigen moments has been proposed by[7]. There are significant improvements in terms accuracy and noise robustness. The invariance properties of radial Tchebichef of moments and their application to pattern recognition tasks have been investigated by R.Mukundan[8].

In this paper, the invariance properties of moments have been used for recognition of objects with different grades of occlusions. Invariant geometric moments and some lower order moments have been utilized for recognition. Invariant moments for different amount of occlusion of a given object is computed. The same moments of the unoccluded object are computed. This above set form a input and target set for learning of a feedforward ANN. The network is trained by Back propagation (BP) algorithm. After training, the moments corresponding to occlusion of other parts of the same object have been used by the trained neural network to predict the moments. The moments, thus predicted have been compared with a previously stored original set of moments. If the predicted error is within a threshold, then object is recognised. The scheme has been succesfully tested for different indoor as well as outdoor images.

II. Problem Statement

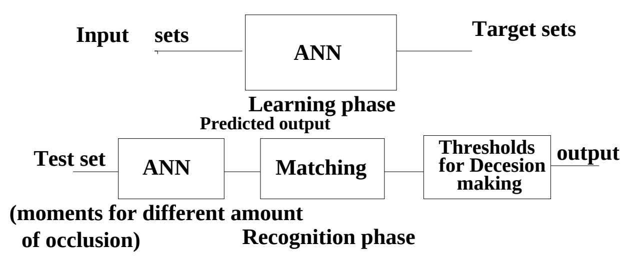

Fig. 1. Block diagram of recognition.

We addressed the issue of object recognition with occlusion of different parts of the object. Since the concern is to recognise the object, some lower order moments that are viewed to represent the smooth part and less edges are considered. In case of machine vision system the object may have translation, rotation and scalling. Geometric moments possess invariance properties to scale, translation and rotation. In this work, we consider some translation invariants moments together with lower order moments for recognition in case of machine vision systems. When an object is occluded, the moments will change. The following Schemes is adopted for such a process. The inputs consists of translational invariants moments and some lower order moments.

Fig.1 shows the block diagram representaion of the scheme. The first block represents the Learning phase.

In this phase, we have considered ten moments upto third order. Moments set with different amounts of occlusion of the object are computed. These set of moments from the inputs training set for a Artificial Neural Network. The target set consists of the moments of the unoccluded objects. The ANN is a feedforward multilayered network. The ANN is trained with these training set using Back-propagation (BP) learning algorithm. Once the network is trained with the data set, the weights are fixed and used for the prediction. Different set of moments for a wide variety of occlusions have been computed and are used as the input sets. The target set is stored and predicted moments are matched with the store moments for decesion making. The output is that either the systems recognises or fails. Particularly, we have investigated with specific situation that is aeroplane flying or indoor object. However, this Scheme can be extended for wide class of objects whose moments are within a certain threshold.

III. Neural Network

The manner in which the neurons of a neural network are structured is intimately linked with a learning used to train the network[9]. Fig.2 presents the architectural layout of a multilayer perceptron.

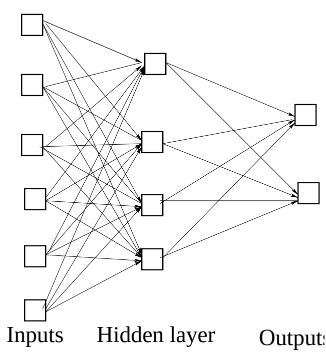

Fig. 2. Schamatic diagram of typical a multilayer feedforward network

This class of a feedforward neural network distinguises by its self by the presence of one or more hidden layers, whose computation nodes are correspondingly called hidden neurons or hidden units. The function of hidden neurons is to intervene between the external input and the network output in some useful manner. By adding one or more hidden layers, the network is enabled to extract higher-order statistics. The ability of hidden neurons to extract higher-order statistics is particularly valuable when the size of the input layer is large. The backpropagation [9] algorithm is used to train a given feed-forward multilayer neural network for a given set of input patterns with known classifications. When each entry of the sample set is presented to the network, the network examines its output response to the sample input pattern. The output response is then compared to the known and desired output and the error value is calculated. Based on the error, the connection weights are adjusted. The backpropagation algorithm is based on Widrow-Hoff delta learning rule in which the weight adjustment is done through mean square error of the output response to the sample input. The set of these sample patterns are repeatedly presented to the network until the error value is minimized. The sequential updating of weights is the preffered method for offline implementation of the back propagation algorithm. For this mode of operation, the salient steps of algorithm can be summarized as follows:

1.Initialization of weights.
2.Presentations of training examples.
3.Forward Computation.
4.Backward Computation.
5.Repeat until the output error is within or preselected threshold.

IV. Moments

A general definition of moment functions ϕ_{pq} of order $(p+q)$ of an image intensity function $f(x,y)$ can be given as follows.

$$\phi_{pq} = \int_x \int_y \psi_{pq}(x,y) f(x,y) dxdy \qquad (1)$$

Here we assume that the real image intensity function $f(x,y)$ is a piece wise contineous function and has bounded support. Where $\psi_{pq}(x,y)$ is the moment weighting kernel. The basis functions may have a range of useful properties that may be passed on to the moments, producing descriptions which can be invariant under rotation, scale and translation. To apply this to digital image, equition (1) need to be expressed in discrete form.

$$\phi_{pq} = \sum_x \sum_y \psi_{pq}(x,y) f(x,y) \qquad (2)$$

Moreover, the orthogonality property of the basis function is passed on to the moments. Thus, non orthogonla basis functions result in non orthogonal moments and orthogonal basis functions result in orthogonal moments. Again orthogonal moments can be divided in to two parts i.e. contineous orthogonal moments and discrete orthogonal moments.

A. Geometric Moments

Geometric moments or regular moments[1] are the most popular types of moments and have been frequently used for a number of image processing tasks. Given the intensity function of an image $f(x,y)$, which is assumed to be piecewise contineous and with compact support, one can define the two dimentional geometric moments of order $(p+q)$ can be defined as

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x,y) dxdy, \qquad p,q = 0,1,2,...... \qquad (3)$$

The two-dimentional moment for a $(N \times N)$ discrete image is given by

$$m_{pq} = \sum_{-\infty}^{+\infty} \sum_{-\infty}^{+\infty} x^p y^q f(x,y) \qquad (4)$$

Moments of low orders can be used for normalizing the density distrubution function $f(x,y)$. The zero order moment m_{00} represents the total image power and can be used, in conjuction with the first order moments m_{10} and m_{01} to locate the centroid of the density distribution, which is given by

$$\overline{x} = \frac{m_{10}}{m_{00}} \quad and \quad \overline{y} = \frac{m_{01}}{m_{00}} \tag{5}$$

The second order moments m_{20}, m_{02}, m_{11} characterise the size and orientation of image. So, the low order moments contain the most basic information regarding the shape, size, and orientation of the image.

V. Results and Discussion

In simulation, we have considered one indoor and two outdoor images. The first outdoor image considered is a car image with a background as shown in Fig.3. Fig.3(a) shows the unoccluded image and the rest of the Fig.3 show the images with different amount of occlusions. The ten lower order moments upto 3rd order are computed for the occluded images as well as the unoccluded ones. The moments of the unoccluded ones form the target set and the moments for the occluded ones form the input. These two constitute the training set. The ANN with ten input nodes, five hidden nodes and ten output nodes is trained using the training set. The parameters used for ANN are learning rate $\eta = 1.5$ slope of the activation function $\lambda = 0.4$. The training examples are presented to the ANN for 550 times and the network learns after 200 iterations. The ANN, thus trained, is used for prediction of moments corresponding to the images as shown in Fig.4. The prediction sets are presented in Table 1 and 2. It is seen from Table 1 and 2 that the prediction error for Fig.4(a), Fig.4(b), Fig.4(c) are less than one percent while for a different object the error for most of the moments are above it. Hence, image of Fig.4(d) is not recognized as a car. The next outdoor image considered is an aeroplane flying. The network is trained with different amount of occlusions. The topology of ANN remains same and the parameters used are $\eta = 1$ and $\lambda = 0.45$. The network, after training is used to predict using the the moments corresponding to different figures shown in Fig.5. For Fig.5(a) and Fig.5(b) the prediction errors are very less (below 0.5 percentage) where as for Fig.5(c) and Fig.5(d) are more than 50 percentage thus, discarding these two objects as aeroplanes. This is observed from Table 4. Similar observations are also made for an indoor images as shown in Fig.6. The same ANN topology with parameters $\eta = 1.5$ and $\lambda = 0.8$ are used for training and after training the images shown in Fig.6 are considered for prediction. It is observed that the prediction error for Fig.6(a) and Fig.6(c) are less while for Fig.6(b) and Fig.6(d) the prediction errors are very high. Thus the scheme recognizes Fig.6(a) and Fig.6(c) as bottle while discards other two images. This high error for Fig.6(b) could be attributed to the position of occlusion of the bottle. Thus, the scheme could be successfully tested for indoor as well as outdoor images.

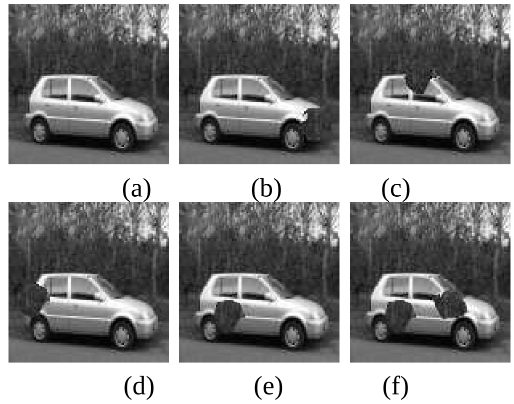

Fig. 3. Training set of images.

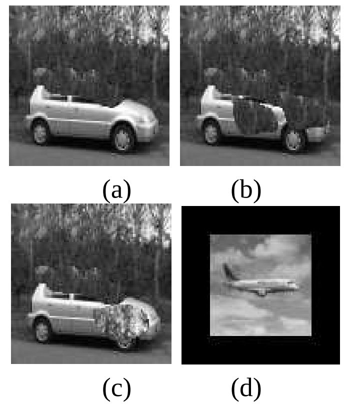

Fig. 4. Test set of images.

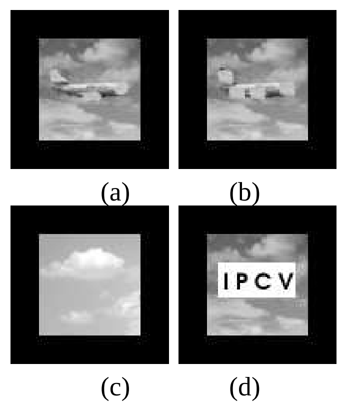

Fig. 5. Test set of images.

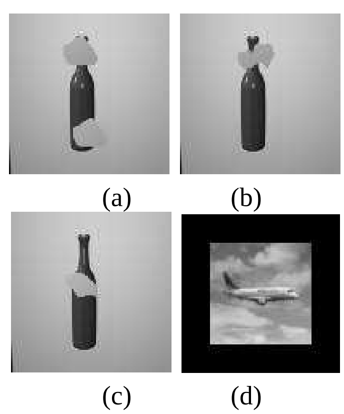

Fig. 6. Test set of images.

TABLE I
PREDICTION AND PERCENTAGE OF ERROR OF TEST IMAGES

μ	Target	Fig.4(a)		Fig.4(b)	
		Prediction	Err(%)	Prediction	Err(%)
μ_{00}	0.000224756	0.000226	0.89	0.000227	1.32
μ_{10}	0.011207859	0.011228	0.18	0.011257	0.44
μ_{01}	0.01163371	0.011654	0.18	0.011683	0.42
μ_{20}	0.194294363	0.194229	0.03	0.194337	0.02
μ_{02}	0.176668947	0.176616	0.02	0.176731	0.03
μ_{11}	0.98896782	0.988948	0.00	0.988919	0.00
μ_{30}	0.073077202	0.073097	0.02	0.073191	0.15
μ_{12}	0.945358454	0.945330	0.00	0.945248	0.01
μ_{21}	0.380255451	0.380195	0.01	0.380242	0.00
μ_{03}	0.302836943	0.302743	0.03	0.302825	0.00

TABLE II
PREDICTION AND PERCENTAGE OF ERROR OF TEST IMAGES

μ	Target	Fig.4(c)		Fig.4(d)	
		Prediction	Err(%)	Prediction	Err(%)
μ_{00}	0.000224756	0.000225	0.44	0.000233	4.01
μ_{10}	0.011207859	0.011211	0.03	0.011413	1.81
μ_{01}	0.01163371	0.011637	0.03	0.011843	1.83
μ_{20}	0.194294363	0.194168	0.06	0.194914	0.31
μ_{02}	0.176668947	0.176556	0.06	0.177314	0.36
μ_{11}	0.98896782	0.988965	0.00	0.988765	0.02
μ_{30}	0.073077202	0.073043	0.04	0.073700	0.85
μ_{12}	0.945358454	0.945376	0.00	0.944814	0.05
μ_{21}	0.380255451	0.380168	0.02	0.380498	0.06
μ_{03}	0.302836943	0.302695	0.04	0.303271	0.14

TABLE III
PERCENTAGE OF ERROR OF TEST IMAGES

Moments	Fig.6(a)	Fig.6(b)	Fig.6(c)	Fig.6(d)
μ_{00}	0.25	18	0.00	44.16
μ_{10}	0.12	9.04	0.01	18.91
μ_{01}	0.10	2.53	0.01	17.98
μ_{20}	0.00	2.66	0.00	1.61
μ_{02}	0.00	0.80	0.00	1.43
μ_{11}	0.15	0.00	0.00	35.37
μ_{30}	0.09	9.00	0.01	18.71
μ_{12}	0.00	0.36	0.00	0.66
μ_{21}	0.00	0.10	0.00	0.21
μ_{03}	0.00	0.41	0.00	0.76

VI. Conclusion

A moment based object recognition system is proposed. In a given scene, an object may be occluded due to various reasons and hence recognition of original object becomes nescessary for machine vision systems. The invariant moments as well as some of the lower order moments are used to form the training set for ANN. Once trained, the ANN can be used for the recognition phase. This proposed scheme is for a fixed enviroment. However, the scheme can also modified for different enviroments. The scheme could be tested sucessfully for indoor as well as outdoor images. Currently work focuses on training as well as reconstruction of objects using Zernike and Pseudo-Zernike moments.

TABLE IV
PERCENTAGE OF ERROR OF TEST IMAGES

Moments	Fig.5(a)	Fig.5(b)	Fig.5(c)	Fig.5(d)
μ_{00}	0.00	0.00	262.32	53.42
μ_{10}	0.02	0.02	99.81	26.05
μ_{01}	0.02	0.02	97.50	25.53
μ_{20}	0.00	0.01	45.26	13.34
μ_{02}	0.01	0.01	44.59	13.09
μ_{11}	0.00	0.00	0.00	0.00
μ_{30}	0.00	0.00	1.64	0.46
μ_{12}	0.01	0.01	50.66	14.75
μ_{21}	0.01	0.01	45.94	13.52
μ_{03}	0.00	0.00	36.45	11.03

Acknowledgment

This work is supported by the MHRD funded Research and Development project on Real time signal and image processing using Genetic Algorithm.

References

[1] Hu.M.K,"Visual Pattern recognition By Moment Invariants, *IRE Transactions on Information Theory*, vol. IT-8, pp. 179-187, Feb. 1962.

[2] Reiss.T.H,"The Revised Fundamental Theorem of Moment Invariants, *IEEE Transactions on Pattern Analysis and Machine Intelligence*,vol. PAMI-13, no. 8, pp. 830-834, Aug. 1991.

[3] Raveendran.P,Jegannathan.P and Omatu,S, "Classifi cation of Elongated and Contracted Images using New Regular Moments, *In IEEE World Congress on Computation*,Orlando, FL, June 26-July 2, pp.4154-4158, 1994.

[4] Tianxu.Z and Jin.L, "Blured Image Recognition based on Moment Invariants, *International Conference on Image Processing*,ICPR-04, pp.2131-2134, 2004.

[5] Teh.C and Chin.R.T, "On Image Analysis by the methods of moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-10, no. 4, pp. 496-513, July 1988.

[6] Khotanzad.A and Hong.Y.H,"Invariant Image Recognition by Zernike Moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-12, no. 5, pp. 489-497, May 1990.

[7] Yap.P.T,Raveendran.P, "Eigenmoments, *Pattern Recognition*, vol.40, pp.1234-1244,2007.

[8] Mukundan.R, "Radial Tchebichef Invariants for Pattern Recognition, *IEEE Conference on image processing*,TENCON-2005.

[9] Haykin,Simon, "*Neural Networks*, Pearson Education Inc.,2004