# SHAPE DISCRIMINATION BY DESCRIPTORS AND MOMENTS USING NEURAL NETWORK

Kyoung-Ok Kim*, Young-Kyu Yang*, Yong-Hui Park**, Tae-Kyun Kim**

Eoeun-Dong, Yusung-ku, Taejeon, Korea 305-333
*Systems Engineering Research Institute/ Korea Institute of Science and Technology
**Dept. of Computer Eng., Chungnam National Univ. Goong-Dong, Yusung-Ku, Taejeon, Korea

Tel.: +82-42- 869-1401 Fax: +82-42- 869-1479
E-mail: ykyang@seri.re.kr, kokim@seri.re.kr

Commission III, Working Group 3

**KEY WORDS : Vision, Training, Feature, Neural**

## ABSTRACT

An important problem in target recognition is the automatic discrimination of the object in a scene regardless of its position, size, and orientation. Object recognition is processed by feature extraction and similarity measurement. This paper is to recognize target objects using moments and Fourier descriptors. The Fourier descriptors and moment features are used as input vectors to the neural network classifier. The difference between the features is that the former deals with contours, while the latter deals with area. This paper presents preprocessing technique and the performance comparison of Zernike moment, Hu's moment invariant and Fourier descriptors as features. Noise is another important factor to affect the recognition accuracy. The contour smoothing as preprocessing for Fourier descriptor is adopted for noise removal.

## INTRODUCTION

Several methods have been studied for object recognition in computer vision and pattern recognition. The process of feature extraction is a very important step in object recognition.

The current approaches to invariant 2D shape recognition include extraction of global image information using regular moments, boundary-base analysis via Fourier descriptors, autoregression models, image representation by circular harmonic expansion, syntactic, and neural network approaches (Kotanzard, 1990). But the global approaches doesn't work very well for occluded objects, so local features should be considered for partially occluded objects. Moments have been utilized as object feature in a number of application for this purpose. The Hu's seven nonlinear functions (Hu, 1962) defined on regular moments are one of the popular type of moments. (Dudant, 1977). But the basis set is not orthogonal. Recently Zernike moment which is known to have strong class separability power is getting popular to derive feature vectors. (Kim, 1994). Moreover Zernike moments used in this study are a class of orthogonal moments. Fourier descriptors which are different from area based moments extract contour features.

Advantages, disadvantages and performance comparison to Fourier descriptors, Zernike moments and moment invariants are also discussed in this paper.

## FEATURE VECTORS

### Moments invariants

(p+q)th moment is defined as

$$m_{pq} = \sum\sum x^p y^q f(x, y) \qquad (1)$$

Central moments can be normalized to become invariant to scale change by defining

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}, \quad \gamma = \frac{p+q}{2} + 1 \qquad (2)$$

And can be invariant to translation by central moment as

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} (x-\bar{x})^p (y-\bar{y})^q f(x,y) \quad (3)$$

$$\bar{x} = \frac{m_{10}}{m_{00}} \text{ and } \bar{y} = \frac{m_{01}}{m_{00}}$$

A set of nonlinear functions defined on $\eta_{pq}$ and invariant to rotation, translation, and scale change have been derived. They are

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{30})^2$$

$$\phi_4 = (\eta_{30} - \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{03} + \eta_{21})$$

$$\phi_7 = (3\eta_{21} - \eta_{30})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

**Zernike Moment**

After Hu's moment was adopted as features for object recognition, several moment, such as Zernike, complex, and Legendre moments were developed. Zernike introduced a set of complex polynomials which form a complete orthogonal set over the interior of unit circle, i, e. $x^2+y^2=1$. Let the set of these polynomials be denoted by $\{V_{nm}(x, y)\}$. The form of these polynomials is

$$V_{nm}(x,y) = V_{nm}(\rho,\theta) = R_{nm}(\rho)\exp(jm\theta), \quad (4)$$

These polynomials has orthogonal basis function (Kotonzard, 1990). Zernike moments are the projection of the image function onto these orthogonal basis functions. The Zernike moment of order n with repetition m for image function, f(x, y) that vanishes outside the unit circle for digital image is

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x,y) V_{nm}^*(\rho,\theta), \quad x^2+y^2 \leq 1. \quad (5)$$

The center of a image is taken as the origin for translation invariance. So ote that the Zernike moment of the rotated image is

$$A'_{nm} = A_{nm}\exp(-jm\phi) \quad (6)$$

Thus Zernike moment can be taken as a rotation-invariant feature. Scale invariance is accomplished by enlarging or reducing each object such that its Zernike-order moment, $m_{oo}$, is set equal to a predetermined value $\beta$.

**Fourier Descriptors**

Fourier descriptors are different from moments in several ways. Fourier descriptor of contour C is defined to be the complex Fourier series expansion of z(t).

$$z(t) = \sum_{n=-\infty}^{\infty} A(n)e^{jnt} \quad (7)$$

where

$$A(n) = \int_0^{2\pi} z(t)e^{-jnt}dt \quad (8)$$

The original sample contour is assumed to be given as a chain code. This chain code representation is converted into a sequence of x-y coordinates. The frequency domain operations which affect the position, size, orientation, and starting point of the contour follow directly from properties of the DFT (Wallace, 1980). To translate a contour a distance x horizontally and y vertically, add x + jy to A(0). To scale a contour by the factor K, multiply each component A(n) of the FD by k. Rotating the contour in the time domain requires multiplying each coordinate by $e^{j\theta}$ where $\theta$ is the angle of rotation. Set A(0) equal to zero to normalize position and scale normalization is accomplished by dividing each coefficient by the magnitude of A(1).

## METHODOLOGY

The process has three stages : preprocessing, feature extraction and learning or recognition. The overall process is shown in Figure 1. At first, the morphological as well as the spatial median filter was applied to suppress small size (one or two pixels) random noise. After preprocessing, an image is segmented by Otsu's thresholding. The binary image is than labeled by using connected component labeling method. With interactive mode, these labeled regions were selected to indicate the training samples which would be used later in training. Then the labeled image is used for training to get feature values. For the feature extraction, three types of features; Hu's moment invariants, Zernike moments and Fourier descriptors were applied. The features of those labeled regions are calculated and inputted to train the Multi-Layer Perceptron. The objects selected in this classifier are normalized with respect to scaling and translation. And also normalized Fourier descriptors of objects were calculated as boundary description. Fourier descriptors are obtained by calculating the FFT value of sampled points. Especially the number of sampling points should be power of 2 to be applied with FFT algorithm. For the recognition phase the Multi-Layer Perceptron (MLP) (George, 1992) which is trained by back propagation learning algorithm was used.

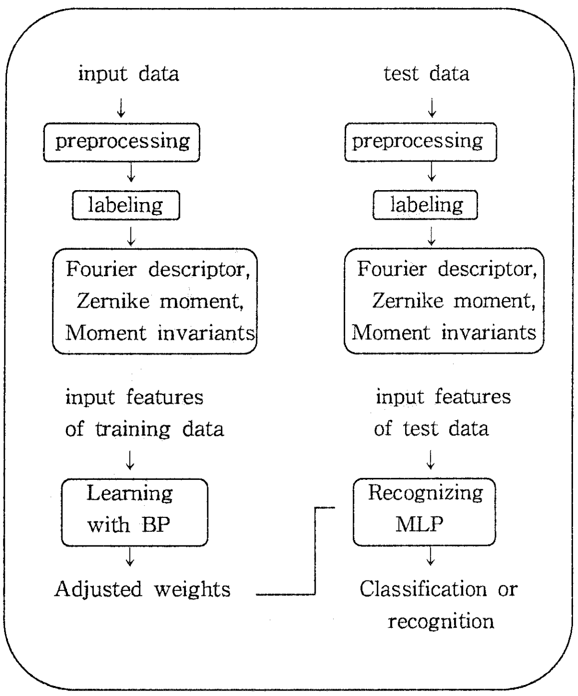



Figure1. Object recognition process

## EXPERIMENTAL STUDY

The classification power of the Zernike moment feature, moment invariants and Fourier descriptor feature selection methods were tested and the results are shown in Figure 2, 3, 6(b). In this experiment three different data sets were collected. The first data set consists of airplanes in JFK airport. For this Figure 4 image, different set with several noise level of random noise and with different orientation is generated. The second data set, Figure 5, is scanned image with 6 toy airplanes, and the third data set, Figure 6(a), is aerial photograph of Seoul Area. The first set of experiment was performed to compare the three features, without preprocessing and with preprocessing. The rotated, translated image of the original scenes and images of other area are used to test the performance of the three feature extraction methods. The multilayer(input, hidden, output layer) neural network is trained with 20 airplanes for the first data set with back propagation algorithm and the parameters used in this experiment is as follows :
iteration 50000, error rate 0.001, learning rate 0.3 and momentum factor 0.7. The same steps were repeated 50 times to enhance the reliability. Error convergence with the morphological filtering converges much faster than the training of original images. For calculation of Fourier descriptor, the outlines were traced yielding chain codes and the normalized Fourier descriptors were computed from all 256 chain codes. The result of second set is shown in Figure 3. This shows that the separability of Zernike moment is superior. The third set of experiment is to detect a certain type of building in the aerial photograph of Seoul area. Figure 6(a) is the image for training. The detected building is shown in Figure 6(b). In this experiment only Zernike moment feature method detects the building.

## CONCLUSION

The discrimination power of the Zernike moment, moment invariant and Fourier descriptor features were tested by a series of experiment on three different data sets using neural network. In general the superiority of Zernike moment features over moment invariants and Fourier descriptors is shown in the results. Especially the obtained classification accuracy of Zernike moment for 6 toy airplane data set is almost perfect. Moreover Zernike moment has lower rate of false detection. The performance of both Zernike and geometrical moment method with preprocessing is better than those of without preprocessing. Zernike moment method exhibits the performance superior to that of moment invariants in terms of false detection. It may be said that Zernike moment method with preprocessing performs much better than the moment based method with the complex and noisy images. On the other hand, Fourier descriptor is much sensitive than moment methods. So the contour smoothing as preprocessing for Fourier descriptor is quite necessary. Thus we can conclude that the Zernike moment features are quite effective for the image classification problem.

## REFERENCES


Dudant, S.A. K. J. Breeding, and R. B. McGhee, 1977. Aircraft identification by moment invariants. IEEE Trans. on Computers, vol. C-26, pp.39-46.

George N, Bebis. and George M. Papadourakis, 1992. Object Recognition Using Invariant Object Boundary Representations and Neural Network

Models, Pattern Recognition, vol. 25, No. 1, pp. 25-44.

Hu, M. K. Visual Pattern Recognition by Moment Invariants, 1962. IRE Trans. Info. Theory, vol. IT-8, pp.179-187.

Khotanzad, A. and Y. Hong, 1990. Rotation Invariant Image Recognition Using Features Selected via a Systematic Method, Pattern Recognition, vol 23 No. 10, pp 1089-1101.

Khotanzad, A. and J. Lu, 1990. Classification of Invariant by Zernike Moments, IEEE Transactions on Patterns Analysis and Machine Intelligence, Vol. 12. No. 5.

Kim, W. and Po Yuan, 1994. A Practical Pattern Recognition System for Translation, Scale and Rotation Invariance, IEEE, Conference on Computer Vision and Pattern Recognition.

Park, Cheol Hoon and Lae Jeong Park, 1993. Invariant Visual Information Processing with Morphological Higher Order Neural Networks, KITE Journal of Electronics Engineering, vol 4, No. 1A.

Wallace, Timothy P. 1980. An Efficient Three-Dimensional Aircraft Recognition Algorithm Using Normalized Fourier Descriptors, Computer Graphics and Image Processing, 13, 99-126

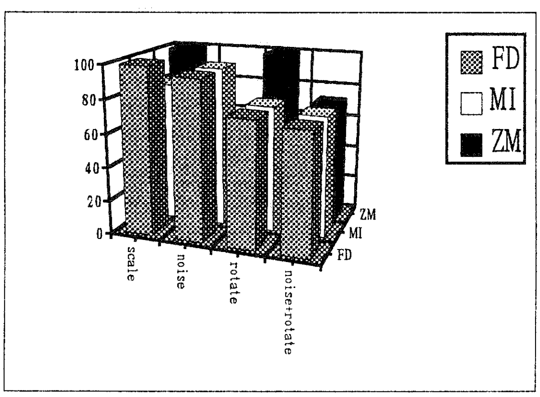
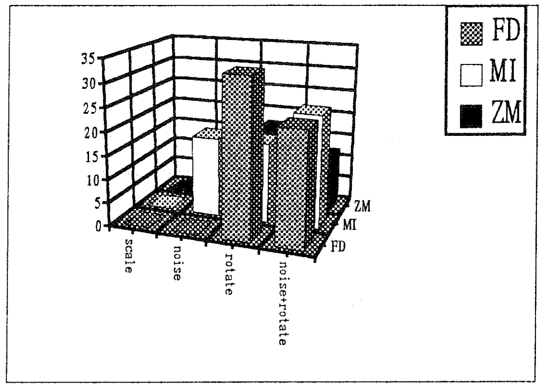
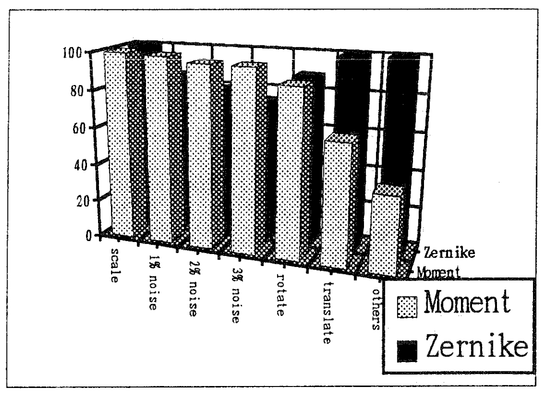


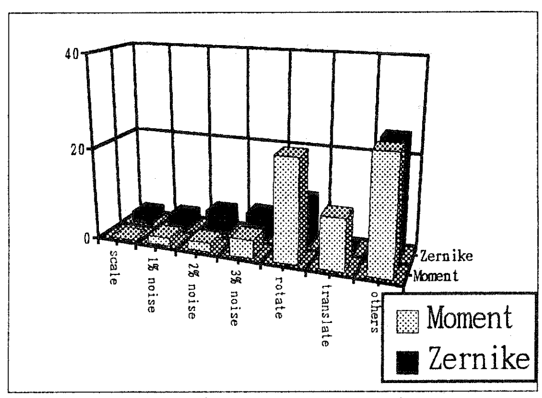

Figure 2. (a) Recognition




(b) False Recognition ( % )




(c) Correct Recognition of airplanes without preprocessing




(d) False Recognition of airplanes with preprocessing ( % )

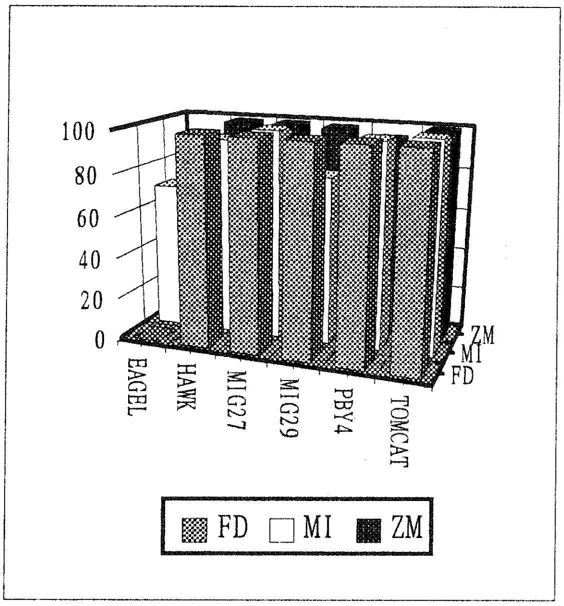



Figure 3. Classfication of six different airplanes

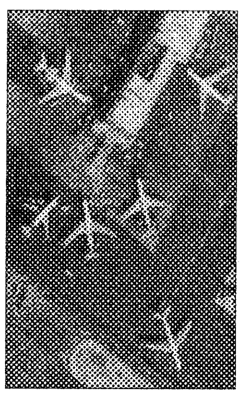


Figure 4. JFK airport scene

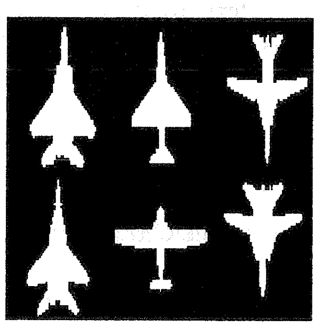


Figure 5. Six different airplanes

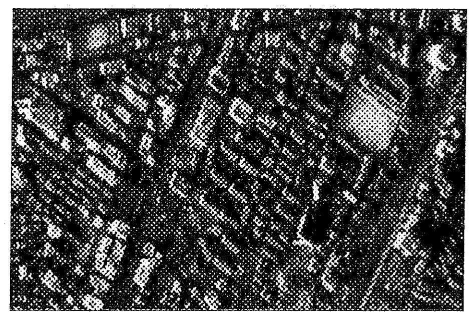


Figure 6. (a) Training image for building detection

(b) The result of building detection