Invariant Pattern Recognition by Zernike Moments using a Neural Network

Guillermo Cámara Chávez
Institute of Mathematics and Computer Science
University of São Paulo
Av. Trabalhador São-Carlense 400, CEP 13560-970
São Carlos, SP - Brazil
e-mail: gcamarac@icmc.sc.usp.br

and

Zhao Liang
Institute of Mathematics and Computer Science
University of São Paulo
e-mail: zhao@icmc.sc.usp.br

Abstract
Pattern recognition has provoked a great interest in the last decades due to the use of the computers. As a consequence, numerous engineering applications have been developed. The complexity of a pattern recognition system is high because much of the information available in the real life is presented in the form of complex patterns, suffering from linear transformations even nonlinear deformations. The objective of this paper is to develop a model for invariant pattern recognition by combining Zernike moments and Fuzzy ART. The former works as invariant feature extractor while the letter acts as a robust classifier. The considered model will efficiently recognize patterns without taking in consideration the possible variations of position, rotation and scale. The experimental results shows the model can achieve high classification accuracy.

Keywords
Pattern Recognition, Zernike Moments, ART, Fuzzy ART

1 Introduction

Pattern recognition is the study of how machines can observe the environment, learn to distinguish patterns of interest from their background, and reasonable decisions about the categories of the patterns [1]. The use in the last decades of computers and electronic devices impulses the study of pattern recognition technics [2]. Interest in the area of pattern recognition has been renewed due to emerging applications which are not only challenging but also computationally more demanding. These applications include *data mining*, document classification, financial forecasting, organization and retrieval of multimedia databases, and biometrics [1]. Patterns that in the real life usually are presented in different positions, having variations in rotation, scale and/or translation. So the research of invariant pattern recognition is very important for real applications. Here term ”invariant” means a quantity which remains unchanged under a certain transformation [1].

A pattern recognition system usually comprises three main components, namely preprocessing, feature extraction and classification [3]. In the preprocessing stage the input image suffers from a variety of operations such as noise removal, segmentation and image enhancement. Feature extraction aims to represent the image in terms of some quantifiable measurements that may be easily utilized in the classification stage [3], there are many feature extractors like template matching [5], Fourier descriptors [6], invariant moments [8]. In the classification stage the patterns are grouped according to similar characteristics. Some of the classifiers are: neural networks, k-means algorithm.

The theory of algebraic invariants of linear transforms arose in connection with a number of problems in analytic geometry and was first formulates by the work of Cayley and Sylvester in the last century [4]. Based on the work of these two mathematicians, Hu published the first paper on the use of image moments for two-dimensional pattern recognition applications [8]. In Hu’s publications, explicit formulas for invariant functions containing the second and third image moments were presented as examples for the algebraic theory of moment invariants. These low-order moments are seven in total and are known as Hu's moment invariants in the literature [4]. Teague [9] suggested the notion of orthogonal moments to recover a image from moments based on the theory of orthogonal polynomials, and has introduced Zernike moments, which allow independent moment invariants to be constructed easily to an arbitrarily high order [10].

Artificial neural networks(ANN) have proven to be powerful classifiers due to the characteristics of learning, fault tolerance and robustness [11]. ART (Artificial Resonance Theory) was developed by Carpenter and Grossberg [12]. These kind of neural network can learn new patterns without forgetting old knowledge. Specifically, the Fuzzy ART [13] has the capability to learn recognition categories rapidly, in response to arbitrary sequences of analog or binary input patterns.

In these paper we present a model for invariant pattern recognition by combining Zernike moments and neural network, Fuzzy ART. The former works as invariant feature extractor while the letter acts as a robust classifier. The considered model will efficiently recognize patterns without taking in consideration the possible variations of position, rotation and scale. The experimental results shows the model can achieve high classification accuracy.

The organization of this paper is as follows. Section II briefly reviews the Zernike moments. In section III we discuss the use of Fuzzy ART as a classifier. Experimental results are presented in section IV. Section V concludes the paper.

2 Zernike Moments

In this section, the Zernike moments descriptor is presented, including the definition and its invariant properties.

2.1 The Definition of Zernike Moments

Zernike introduced a set of complex polynomials which form a complete orthogonal set over the interior of the unit circle, i.e., $x^2 + y^2 = 1$. These polynomials are denoted by $V_{nm}(x,y)$, it could be express in polar coordinates as [14]:

$$V_{nm}(x,y) = V_{nm}(\rho,\theta) = R_{nm}(\rho)exp(jm\theta) \tag{1}$$

where

n	Positive integer or zero.
m	Positive and negative integers subject to constraints $n - \|l\| = par, \|m\| \leq l$.
ρ	Length of vector from origin to (x,y) pixel.
θ	Angle between vector ρ and x axis in counterclockwise direction.
$R_{nm}(\rho)$	Radial polynomial defined as

The Zernike radial ponymonials $R_{nm}(\rho)$ are defined as

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|/2} (-1)^s . \frac{(n-s)!}{s!(\frac{n+|m|}{2}-s)!(\frac{n+|m|}{2}-s)!}\rho^{n-2s} \tag{2}$$

Note that $R_{n,-m}(\rho) = R_{nm}(\rho)$
It can be easily verify that these polynomial are orthogonal [14].

Zernike moments are the projection of the image function onto these orthogonal basis function. The Zernike moment of order n with repetition m for a continuous image function $f(x,y)$ that vanishes outside the unit circle is [14]:

$$Z_{nm} = \frac{n+1}{\pi} \int\int_{x^2+y^2=1} f(x,y)V^*_{nm}(\rho,\theta)dxdy \tag{3}$$

To compute the Zernike moments of a given image, the center of the image is taken as the origin and pixel coordinate are mapped to the range unit circle. The pixels that follow outside the unit circle are not considered in the computation [14]. So images transformed by this procedure are already translation and scaling invariant.

2.2 Rotation Invariant Property

Consider a rotation of the image through angle α. Lets denote the rotate image as f^r, the relationship between the original and rotated images in the same polar coordinates is

$$f^r(\rho,\theta) = f(\rho,\theta-\alpha) \tag{4}$$

We can express the Zernike moments in polar coordinates by changing the variables in (3).

$$\int_A \int \partial(x,y)dxdy = \int_G \int \partial[p(\rho,\theta), q(\rho,\theta)]\frac{\delta(x,y)}{\partial(\rho,\theta)}d\rho d\theta \tag{5}$$

where $\frac{\delta(x,y)}{\delta(\rho,\theta)}$ denotes the Jacobian of the transformation and is the determinant of the matrix

$$\frac{\partial(x,y)}{\partial(\rho,\theta)} = \begin{bmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \theta} \end{bmatrix}. \tag{6}$$

Where $x = \rho\cos\theta$ and $y = \rho\sin\theta$, the Jacobian becomes ρ. Hence

$$\begin{aligned} Z_{nm} &= \frac{n+1}{\pi}\int_0^{2\pi}\int_0^1 f(\rho,\theta)V_{nm}^*(\rho,\theta)\rho d\rho d\theta \\ &= \frac{n+1}{\pi}\int_0^{2\pi}\int_0^1 f(\rho,\theta)R_{nm}(\rho)exp(-jm\theta)\rho d\rho d\theta \end{aligned} \tag{7}$$

The Zernike moment of the rotated image in the same coordinate is

$$Z_{nm}^r = \frac{n+1}{\pi}\int_0^{2\pi}\int_0^1 f(\rho,\theta-\alpha)R_{nm}(\rho).exp(-jm\theta)\rho d\rho d\theta \tag{8}$$

By a change of variable $\theta_1 = \theta - \alpha$

$$\begin{aligned} Z_{nm}^r &= \frac{n+1}{\pi}\int_0^{2\pi}\int_0^1 f(\rho,\theta_1)R_{nm}(\rho).exp(-jm(\theta_1+\alpha))\rho d\rho d\theta \\ &= \left[\frac{n+1}{\pi}\int_0^{2\pi}\int_0^1 f(\rho,\theta_1)R_{nm}(\rho).exp(-jm(\theta_1))\rho d\rho d\theta_1\right]exp(-jm\alpha) \\ &= Z_{nm}exp(-jm\alpha) \end{aligned} \tag{9}$$

Equation (9) shows that Zernike moments have simple rotational transformation properties; each Zernike moment merely acquires a phase shift on rotation. This simple property leads to the conclusion that the magnitudes of Zernike moments of a rotated image function remain identical to those before rotation.

3 Fuzzy ART

Fuzzy ART inherits the design features of other ART models, and incorporates computations from fuzzy set theory into the ART1 neural network. As a consequence Fuzzy ART can learn and classify analog patterns.

3.1 Fuzzy ART Algorithm

Figure 1 shows the basic architecture of Fuzzy ART. Each input pattern I is a m-dimensional vector $(I_1, I_2, I_3, \ldots, I_m)$, Each cluster (j) corresponds to a vector $W_j = (W_{j1}, W_{j2}, W_{j3}, \ldots, W_{jm})$ of adaptive weight. The number of possible cluster $n(j = 1, 2, \ldots, n)$ is arbitrary. The Fuzzy ART weight vector is equivalent in one vector to both ART1 weight vectors *bottom-up, top-down*, that is, subsumes both vectors [15]. Each ART system includes a field, F_0, of nodes that represents a current input vector I^T; a field, F_1, that receives both bottom-up input from F_0 and top-down input from a field, F_2, that represents the active code, or cluster (Figure 1). The F_1 activity vector is denoted **x**=$(x_1, \ldots, x_m)$ and the F_2 activity vector is denoted **y**=$(y_1, \ldots, y_n)$. The number of nodes in each field is arbitrary, see Figure 1.

1. *Initialize*

   (a) Initially, each cluster is said to be $uncommitted$, after a category is selected for coding it becomes $committed$, and the weight vector W_j is set as

   $$W_{j1}(0) = W_{j2}(0) = \ldots = W_{jm}(0) = 1 \tag{10}$$

   (b) Then, a choice parameter α, a learning rate β, and a vigilance parameter ρ are set.

   $$\alpha > 0, \beta \in [0,1], \rho \in [0,1].$$

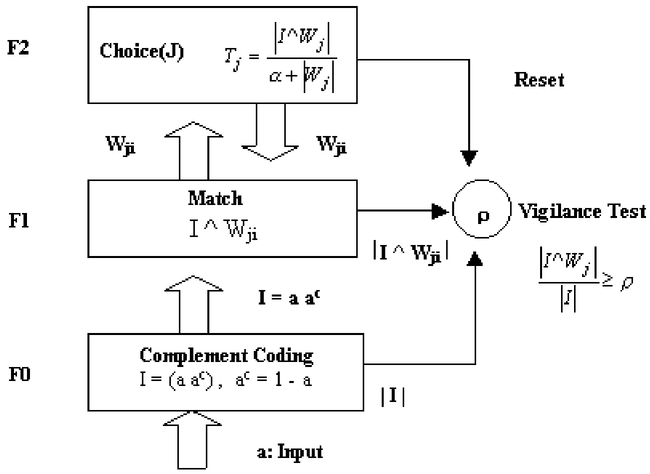


Figure 1: Fuzzy ART

2. *Complement Coding*
   (a) To improve the reliability of category choice, input a is expanded with complement coding as in Eq. 11
   $$I = (a, a^c), a^c = 1 - a \tag{11}$$
3. *Category choice*
   (a) For each I and cluster (F_2 node) j, the choice function T_j is defined by
   $$T_j(I) = \frac{|I \wedge W_j|}{\alpha + |W_j|}, \tag{12}$$
   where the fuzzy AND operator $\wedge$ is defined by $(x \wedge y)_i = min(x_i, y_i)$ and $|\ |$ is defined by
   $$|x| = \sum_{i=1}^{m} |x_i| \tag{13}$$
   (b) The system makes a cluster selection when more than one cluster could be selected at a given time. The index J denotes the chosen cluster, where
   $$T_J = max\{T_j : j = 1, \dots, n\}. \tag{14}$$
   If more than one T_J is maximal, the system chooses the category with the smallest j index. Nodes become $committed$ in order $j = 1, 2, 3, \dots$.
4. *Resonance or Reset*
   (a) The resonance occurs if the match function of the chosen cluster meet the vigilance criterion; where
   $$\frac{|I \wedge W_J|}{|I|} \geq \rho. \tag{15}$$
   The learning processes is done according to the equation
   $$W_J^{(new)} = \beta(I \wedge W_J) + (1 - \beta)W_J^{(old)}. \tag{16}$$
   Fast learning corresponds to $\beta = 1$, which is the learning rule in Eq. 17
   $$W_J^{(new)} = \beta(I \wedge W_J). \tag{17}$$
   (b) *Mismatch reset* occurs if
   $$\frac{|I \wedge W_J|}{|I|} < \rho. \tag{18}$$

Then the value of the choice function T_J is set to 0 for the duration of the input presentation to prevent the persistent selection of the same category selection during search. A new index J is chosen by (14). The search continues until the chosen J satisfies (15). If no one of the clusters is selected, then a new cluster must be incremented in $F2$ field.

3.2 Fast-Commit Slow-Recode Option

For efficient coding of noisy input sets, it is useful to set β = 1 when J is an uncommitted node, and then to take $\beta < 1$ after the category is committed. Then $W_J^{(new)} = I$ the first time category J becomes active [16], this is done for guaranty fast initial learning and slow forgetting rate.

3.3 Input Normalization/ Complement Coding Option

Proliferation of categories is avoided in Fuzzy ART if inputs are normalized; that is, for some $\gamma > 0$,

$$|I| \equiv \gamma \tag{19}$$

for all inputs I. Normalization [16] can be achieved by preprocessing each incoming vector **a**, for example, setting

$$I = \frac{a}{|a|} \tag{20}$$

Complement coding is a normalization rule that preserves amplitude information, solves the category proliferation problem, that can occur when a large number of inputs erode the norm of weight vectors. Complement coding represents both the on-response and the off-response to an input vector **a**. Let **a** represent the on-response, and a^c the off response, where

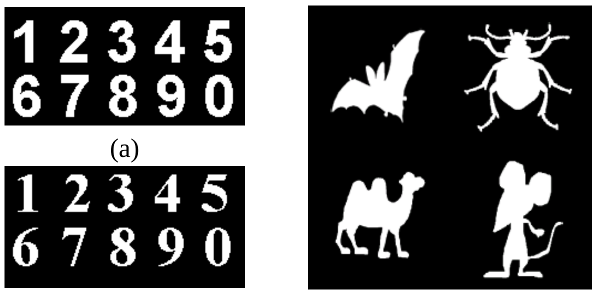
$$a_i^c \equiv 1 - a_i \tag{21}$$

The complement coded input **I** to the field F_1 is the 2m-dimensional vector.

$$I = (\mathbf{a}, \mathbf{a}^c) \equiv (a_1, \ldots, a_m, a_1^c, \ldots, a_m^c) \tag{22}$$

4 Experimental Results

The database consist of 128x128 bits monochromatic images representing the digits from 0 to 9 in *Arial* and *Times new roman* font, Figure 2(a), and four animals images, Figure 2(b).



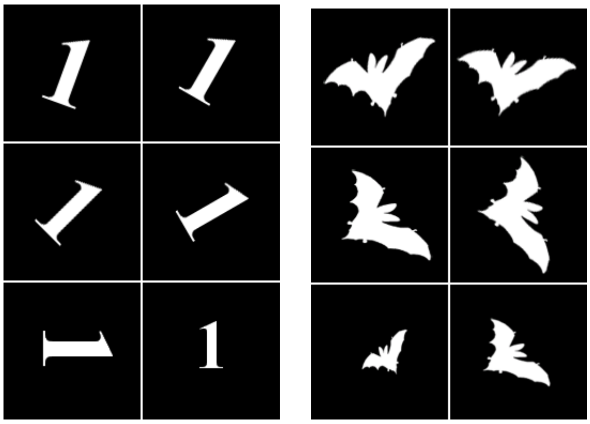
(b) (c)

Figure 2: (a) and (b) First database consists of patterns representing digits from 0 to 9 in *arial* and *times new roman* font. (b) Second database is formed for four different patterns of animals.

In the first data set six different images were generated for each digit considering different orientation, scale and translation. In Figure 3 we can see some variations generated from the digit 1. Following the same process, 56 different images were generated for each digit. The whole numerical database has 560 patterns. In the second data set similar geometric modifications occurred with the images of animals, that's it, were rotated, translated and/or scaled. In the animals database also six different images were generated. The modifications done to every pattern in animal data set were as follow: rotation angles were $20°, 45°, 90°, 120°$, scale and rotated, 0.5 scale. The whole animal data set consist of 576 patterns.

In numerical (*Arial* font) and animal data set not only have been applied geometrical transformations also we increment noisy patterns, Figure 4. Patterns with salt and pepper noise were used for training. The noise density were of 0.02, 0.05, and 0.1. We only apply noise to *Arial* font and animal data set. *Times new roman* font only suffered rotational, scale and translation modifications. The objective for leaving *Times new roman* font noiseless is to compare classification results with noisy data sets.

Figure 5 shows 7 different feature vectors obtained from the images showed in Figure 3 (a) through the Zernike moments. All vectors are very similar, demonstrating graphically the invariance in rotation, translation and scale of Zernike moments. It is important to notice that Zernike moments are only rotational invariant. Translation and scaling invariance is achieved using regular moments [14]. In this work regular moments was used only for supplying the



(a) (b)

Figure 3: (a) Geometrical deformations of pattern “1”. The rotated images of pattern “1” and a scale version. From top left to right, rotation angles are $20°, 30°, 45°, 60°, 90°$ and a 0,5 scale version. (b) Scale, rotation and translate transformations of animal pattern.From top left to right, rotation angles are $20°, 45°, 90°, 120°$, 0.5 scale, scale and rotated.

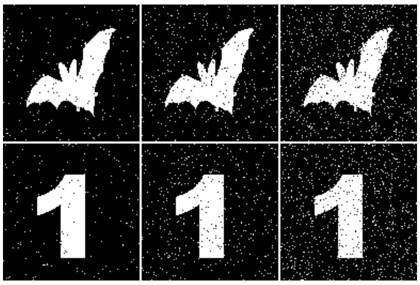


Figure 4: One example of noisy versions of patterns “1” and “bat” with 0.02, 0.05 and 0.1 noise density.

centroid of every pattern. As we see in Figure 5 experiments based on Zernike moments vector is well carried out for scale change up to 50%. For bigger scale modifications it is necessary to normalize (scale and translation) the image using regular moments [9].

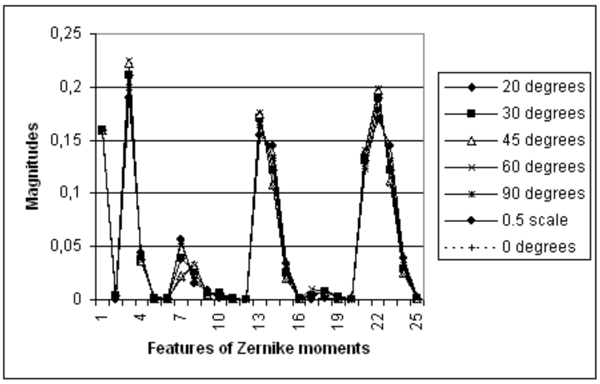


Figure 5: All Zernike moments of the seven different silhouettes.

On the other hand Figure 6 shows three feature vectors extracted from three different patterns “0”, “1” and “2” shown in Figure 2(b) using the 8th-order (25 features) Zernike moments. The difference among those vectors is visible allowing to classify them in different groups.

It is important to note that low order moments are similar for different patterns, even though they belongs to different classes. On the other hand high order moments are able to distinguish different patterns, but the computational cost is high too.

In this study we work with 8th-order moments having 25 features. It may be noted that the total number of moment

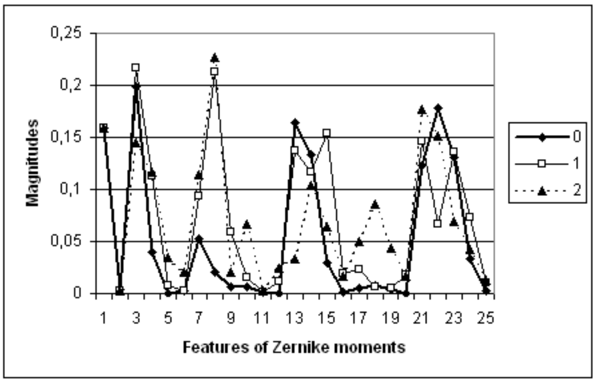


Figure 6: 8th-order Zernike moments of patterns “0” through “2”.

	Times	Arial	Animals
Choice parameter (α)	40	40	0.1
Learning rate (β)	1	1	1
Vigilance parameter (ρ)	0.95	0.94	0.934

Table 1: Parameters values.

terms from zero order up to order *n*, is

$$\Omega = \frac{(n+3)(n+1)}{4} \quad \text{n is odd} \tag{23}$$

$$\Omega = \frac{(n+2)(n+2)}{4} \quad \text{n is even.} \tag{24}$$

The selected parameters for experiments are described in the Table 1 The choice and learning rate parameters are identical in numerical data base. As Fuzzy ART uses a non-supervised training algorithm, the clusters have to be detected by the network. The Fuzzy ART grouped pattern with similar characteristics, as we see in the Table 2. The patters representing “6” and “9” digits present similar Zernike moments, consequently, they were grouped in the same cluster. Figure 7 describes graphically the features vectors of patterns “6” and “9”, both of them are very similar. Those two patterns are similar because the digit *six* can be obtained rotating the number “9” 180 degrees. Even for the human eye, it is difficult to notice the difference. We need a mark in the roof of the background to difference both of them “$\underline{6}$” and “$\underline{9}$”.

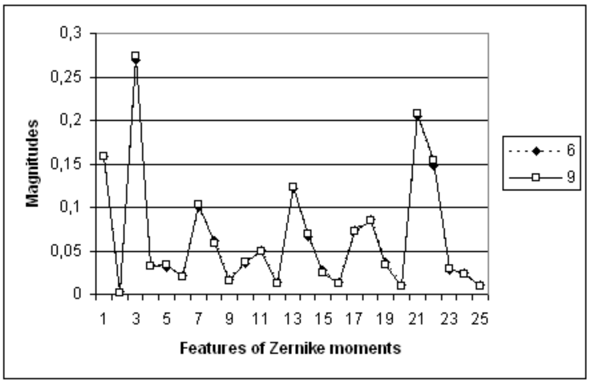


Figure 7: Plotting Zernike moments of patterns “6” and “9”.

The performance of the 8th-order Zernike moments and the Fuzzy ART classifier on different data sets is shown in Table 3. The classification accuracy achieved in this work was satisfactory, in numerical set 100.00% and 86.5%, and in animal set 99.65%. Patterns with bigger density noise than we used in this work were classified in different clusters. Reducing the ratio when the Zernike moments were extracted reduce the influence of noise in the image.

	Pattern ”6”	Pattern ”9”
Z_{20}	0.273	0.269
Z_{22}	0.031	0.034
Z_{31}	0.033	0.030
Z_{33}	0.019	0.021

Table 2: Zernike moments of patterns “6” and “9”.

Order Moment	Classification Accuracy %
Times new roman	100%
Arial	86.5%
Animal database	99.65%

Table 3: Classification results of data sets.

5 Conclusions

The proposed model can efficiently recognize patterns regardless of their possible position, rotation and scale variations. This model consists of two systems: feature detection and classification system. The former works as invariant feature extractor while the letter acts as a robust classifier. It is necessary increase the number of Zernike moments (higher orders) when the images are similar they belongs to different classes. Low order moments describe global characteristics and high order patterns details. The obtained classification accuracy for a 10-class digit data set is 100% in *Times new roman* font (noiseless) and 86.5% in *arial* font with noisy patterns, and in 4-class animal data base also with noisy patterns is 99.65%. We can conclude that the proposed features set and the accompanying features selection method are effective for pattern classification problem.

Acknowledgments: The authors thank to CNPq/NSF for financial support.

References

[1] Anil K. Jain, Robert P. W. Duin, and Jianchang Mao, Statistical pattern recognition: A review, *IEEE Trans. Pattern Anal. Machine Intell.*, 22(1), January 2000, 4-37.

[2] F.J. Cortijo Bon, Introducción al Reconocimiento de Formas, $http : //www-etsi2.ugr.es/depar/ccia/rf/www/tema1\_00-01\_www/$, october 2000.

[3] Y. C. Chim, A. A. Kassim, Y. Ibrahim, Character Recognition using Statistical Moments, *Image Vision Computing*, 17, 1999, 299-307.

[4] Yajun Li, Reforming The Theory of Invariant Moments for Pattern Recognition, *Pattern Recognition*, 25(7), 1992, 723-730.

[5] R.J. Schalkoff. *Digital Image Processing and Computer Vision* (New York : Wiley, c1989).

[6] D. Ballard, C. Brown, *Computer Vision* (New Jersey: Prentice-Hall, 1982).

[7] G. H. Granlund, Fourier processing for handwritten character recognition, *IEEE Trans. Comput.*, 21, 1992, 195-201.

[8] M. K. Hu, Visual Pattern Recognition by Moment Invariants, *IRE Trans. Inf. Theory IT-8*, 1962, 179-187.

[9] M. R. Teague, Image Analysis via the General Theory of Moments, *J. Opt. Soc. Amer.*, 70, Aug. 1980, 920-930.

[10] Cho-Huak Teh, On Image Analysis by the Methods of Moments, *IEEE Trans. Pattern Anal. Machine Intell.*, 10(4), July 1988, 496-513.

[11] E. Barnard, D. Casasent, Invariance and Neural Nets, *IEEE Trans. on Neural Network.*, 2(5), September 1991, 498-508.

[12] G. A. Carpenter and S. Grossberg, Amassively parallel architecture for a self-organizing neural pattern recognition machine, *Computer Vision, Graphics, Image Process*, 37, September 1987, 54-115,

[13] G. A. Carpenter, S. Grossberg and D. B. Rosen, Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system, *Neural Network*, 4, September 1991, 759-771,

[14] A. Khotanzand, Y. Hua Hong, Invariant Pattern Recognition by Zernike Moments, *IEEE Trans. Pattern Anal. Machine Intell.*, 12(5), 1990, 489-497.

[15] Mun-Hwa Kim, Dong-Sik Jang, Young-Kyu Yang, A robust-invariant pattern recognition model using Fuzzy ART, *Pattern Recognition*, 34, 2001, 1685-1696.

[16] G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, and D. B. Rosen, Fuzzy ARTMAP: A Neural Network Architecture for incremental Supervised Learning of Analog Multidimensional Maps, *IEEE Trans. on Neural Network*, 3(5), September 1992, 698-712,