FARSI HANDWRITTEN CHARACTER RECOGNITION WITH MOMENT INVARIANTS

Mehdi Dehghan and Karim Faez

Electrical Engineering Department
AmirKabir University of Technology
Hafez Ave., Tehran, Iran, 15914

Email : kfaez@cic.aku.ac.ir

ABSTRACT

This paper introduces an experimental evaluation of the effectiveness of utilizing various moments as pattern features in recognition of the handwritten Farsi characters. The moments that have been used are Zernike moments, Pseudo Zernike moments, and Legendre moments. We have used an unsupervised neural network (ART2) for this application, so that the clusters are formed only based on inherent properties of pattern features. The performance of classification is dependent on the moment order as well as the type of the moment invariant, but the classification error rate was below %10 in all cases. The Pseudo Zernike moments of order of 5 had the best performance among all the moment invariants. Its error rate and discrimination factor were %3.06 and %96.92 respectively.

1. INTRODUCTION

Orthogonal moments that are invariant to the basic image transformation such as translation, scaling, and rotation are of great importance in pattern recognition [3][6]. This paper introduces an experimental evaluation of the effectiveness of utilizing orthogonal moments such as Zernike moments, Pseudo Zernike moments, and Legendre moments in recognition of the handwritten Farsi characters.

Farsi text is cursive and is written from right to left. A Farsi character might have several shapes. The shape of a character depends on its relative position in the word. For example, the character "ع" is written differently in a word ; at the begining of the word as "عـ" in the middle of the word as "ـعـ", at the end of the word as "ـع", and in isolation as "ع". In addition, some Farsi characters have the same shapes and differ from each other only by existing of dots or zigzag bar. The isolated Farsi character set is shown in fig. 1. This work concerns with recognition of the main body of isolated handwritten Farsi characters which consist of 18 shapes. Since the isolated character "م" has another shape ("ـم") and when the character "ا" proceeded by the character "ل" is written as "لا", the total shapes of the main body of isolated Farsi characters increases to 20 shapes as in fig. 2. For each category, 160 samples was gathered from five famous Iranian calligraphers (totally 3200 samples), which half of them used for training and another half for testing the neural network.

The rest of this paper is organized as follows. Section 2 gives a review of calculating orthogonal moments. A summary on Adaptive Resonance Network (ART2) is given in section 3. Section 4 gives the experimental results and conclusion.

2. ORTHOGONAL MOMENTS

Selection of a feature extraction method is one of the most important factors in achieving high recognition performance[3]. In order to recognize variations of a character, invariant features must be used. Therefore we have used Zernike, Pseudo

Zernike, and Legendre moments for this research. Since these moments are calculated using scaled Geometric moments, they are invariant to translation and scaling. Also these moments are robust to high frequency noise, since we never use very high order terms. These moments have the reconstructability property too.

Geometric moments: The Geometric moments of order p+q of a digital image are defined as:

$$M_{pq} = \sum_x \sum_y f(x,y) x^p y^q \quad (1)$$

where $p,q = 0,1,2,\ldots,\infty$.

The translation invariant Central moments are obtained by placing origin at the centroid of the image.

$$\mu_{pq} = \sum_x \sum_y f(x,y)(x-x_0)^p (y-y_0)^q \quad (2)$$

where

$$x_0 = \frac{M_{10}}{M_{00}} \text{ and } y_0 = \frac{M_{01}}{M_{00}}.$$

Then, the scale invariant Central moments are defined as:

$$G_{pq} = \frac{\mu_{pq}}{\alpha^{(p+q+2)/2}} \quad (3)$$

where $\alpha = M_{00}$.

Finally, the scale invariant Radial-Geometric moments are defined as:

$$R_{pq} = \frac{\sum_x \sum_y f(x,y)(\tilde{x}^2+\tilde{y}^2)^{1/2} \tilde{x}^p \tilde{y}^q}{\alpha^{(p+q+3)/2}} \quad (4)$$

where $\tilde{x} = x - x_0$ and $\tilde{y} = y - y_0$.

Zernike moments : Zernike polynomials are an orthogonal set of complex-valued polynomials:

$$V_{nm}(x,y) = R_{nm}(x,y).\exp(jm\tan^{-1}(\frac{y}{x})) \quad (5)$$

where $x^2+y^2 \le 1$, $j=\sqrt{-1}$, $n \ge 0$, $|m| \le n$ and $n-|m|$ is even and Radial polynomials $\{R_{nm}\}$ are defined as:

$$R_{nm}(x,y) = \sum_{s=0}^{\frac{n-|m|}{2}} B_{n|m|s}(x^2+y^2)^{n/2-s} \quad (6)$$

where

$$B_{n|m|s} = \frac{(-1)^s (n-s)!}{s!(\frac{n+|m|}{2}-s)!(\frac{n-|m|}{2}-s)!} \quad (7)$$

The complex Zernike moments of order n and repetition m are given by:

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x,y) V^*_{nm}(x,y) \quad (8)$$

where $x^2+y^2 \le 1$ and symbol * denotes the complex conjugate operator[2].

The Zernike moments can be computed by the scale invariant Central moments as follows:

$$A_{nm} = \frac{n+1}{\pi} \sum_{\substack{k=|m| \\ n-k=even}}^{n} \sum_{a=0}^{b} \sum_{d=0}^{|m|} (-j)^d \cdot \binom{|m|}{d}\binom{b}{a} B_{n|m|s} G_{k-2a-d,2a+d} \quad (9)$$

where

$$b = \frac{n-|m|}{2} - s \text{ and } j = \sqrt{-1}.$$

Pseudo Zernike moments: The Zernike moments in equ. (8) become Pseudo Zernike moments if the Radial polynomials $\{R_{nm}\}$ defined as below are used to compute the polynomials with condition n-|m| is even eliminated [2].

$$R_{nm}(x,y) = \sum_{s=0}^{n-|m|} D_{n|m|s}(x^2+y^2)^{\frac{n-s}{2}} \quad (10)$$

where

$$D_{n|m|s} = \frac{(-1)^s (2n+1-s)!}{s!(n-|m|-s)!(n+|m|+1-s)!} \quad (11)$$

The Pseudo Zernike moments can be computed by the scale invariant Central moments and the Radial-Geometric moments as follows [6]:

$$A_{nm} = \frac{n+1}{\pi} \sum_{\substack{s=0 \\ n-s-m=even}}^{n-|m|} D_{n|m|s} \sum_{a=0}^{k} \sum_{b=0}^{m} (-j)^b \binom{k}{a}\binom{m}{b} B_{n|m|s} G_{2k-2a+m-b,2a+b} + \frac{n+1}{\pi} \sum_{\substack{s=0 \\ n-s-m=odd}}^{n-|m|} D_{n|m|s} \sum_{a=0}^{d} \sum_{b=0}^{m} (-j)^b \binom{d}{a}\binom{m}{b} B_{n|m|s} R_{2d-2a+m-b,2a+b} \quad (12)$$

where $k = \frac{n-s-m}{2}$, $d = \frac{n-s-m-1}{2}$ and $j = \sqrt{-1}$.

Legendre moments: The (n,m) order Legendre moments are defined as:

$$\lambda_{nm} = \int_{-1}^{+1}\int_{-1}^{+1} P_n(x) P_m(y) f(x,y) dxdy \quad (13)$$

where the nth order Legendre polynomial is given by:

$$P_n(x) = \sum_{j=0}^{n} a_{nj} x^j$$
$$= \frac{1}{2^n n!} \cdot \frac{dn}{dx^n}(x^2-1)^n \quad (14)$$

where

$$a_{nj} = \begin{cases} 0 & \text{n-j=odd} \\ \frac{1}{2^j}(-1)^p \binom{n}{p}\binom{2n-2p}{n} & \text{n-j=even} \end{cases} \quad (15)$$

and $p = \frac{n-j}{2}$.

The Legendre moments can be computed by Geometric moments as follows [2]:

$$\lambda_{nm} = \frac{(2n+1)(2m+1)}{4} \sum_{j=0}^{n}\sum_{k=0}^{m} a_{nj} a_{mk} M_{jk} \quad (16)$$

3. Adaptive Resonance Classifier

The ART2 neural network clusters inputs by using unsupervised learning. The basic architecture of an adaptive resonance neural net involves three groups of neurons (fig. 3): An input processing field(F1 layer), the cluster units(F2 layer), and a mechanism to control the degree of similarity of patterns placed on the same cluster(a reset mechanism)[5].

The F1 layer and the F2 layer are connected to each other by bottom-up and top-down weights. The F2 layer is a competitive layer. The cluster unit with the largest net input becomes the candidate to learn the input pattern. Whether or not this cluster unit is allowed to learn the input pattern depends on how similar its top-down weight vector is to the input vector. The degree of similarity is controlled by the vigilance parameter. If the similarity is below the vigilance parameter, the cluster unit will become inhibited and another candidate is chosen. This process continues until an acceptable cluster unit is chosen (so learning will occur) or all cluster units are inhibited(so the vigilance parameter will be decreased and the process will be repeated)[5].

We have used slow learning, in which only one iteration of the weight update equations occurs on each learning trial. In testing phase, the input vectors are variants of the set of vectors used to train the neural net. The classifier is expected to recall the correct category to which each test pattern belongs. If the recalled category of a test pattern is not the desired category, a misclassification has occurred. If all cluster units inhibited by the reset mechanism for a test pattern, that pattern will be rejected. During testing no learning takes place[4].

4. Experimental results

The main body of isolated handwritten Farsi characters consist of 20 classes. We gather 3200 samples from five famous Iranian calligraphers. Half of the samples used for training the net and another half for testing. The moments are calculated up to 8th order for all samples. Feature moments from order of 3 to 8 were used as feature patterns in separate experiments. In each experiment, the maximum number of training epochs was 10000 and the maximum number of cluster units was 20. The learning rate and the vigilance parameter were set to 0.7 and 0.99 respectively. The dimension of input vectors depends on the order and the type of moments. In training phase, all input patterns were grouped to one of the clusters(no rejection). Classification error rate in training phase is shown in fig.4.

In testing phase, a test pattern may be grouped in a correct cluster, or in an incorrect cluster, or may be rejected. In order to have an objective evaluation of the performance of the various moments, we used the following measures[1]:

1) Discrimination factor = 100 NC/(N-NR)
2) Error percentage = 100 NE /N

where

N = total number of test samples (N = 1600 in this study).
NC = number of correctly classified test samples.
NR = number of rejected test samples.
NE = number of misclassified test samples.

Figures 4 to 6 present a comparison between effectiveness of various moments in handwritten Farsi character recognition. As shown in the figures, The error rate is below %10 in all cases. The Pseudo Zernike moments of order of 5 have the best performance among all the moment invariants.

Its error rate and discrimination factor are %3.06 and %96.92 respectively.

REFERENCES

1) S. O. Belkasim, M. Shridhar and M. Ahmadi, "Pattern Recognition With Moment Invariants : A Comparative Study and New Results," Pattern Recognition, Vol. 24, No. 12, PP.1117-1138, 1991.

2) C. H. Teh and R. T. Chin, " On Image Analysis By The Methods Of Moments," IEEE Trans. Patt. Anal. Mach. Intell., Vol. 10, No. 4, PP. 496-513, 1988.

3) O. D. Trier, A. K. Jain and T. Taxt," Feature Extraction Methods For Character Recognition - A Survey," Pattern Recognition, Vol. 29, No. 4, PP. 641-662, 1996.

4) K. W. Gan and K. T. Lua," Chinese Character Classification Using An Adaptive Resonance Network," Pattern Recognition, Vol. 25, No. 8, PP. 877-882, 1992.

5) L. Fausett, *Fundamentals Of Neural Networks*, Prentice-Hall, PP. 218-287, 1994.

6) R. R. Bailey and M. Srinath, "Orthogonal Moment Features for Use With Parametric and Non-Parametric Classifiers," IEEE Trans. Patt. Anal. Mach. Intell., Vol. 18, No. 4, PP. 389-398, 1996.

FIGURES

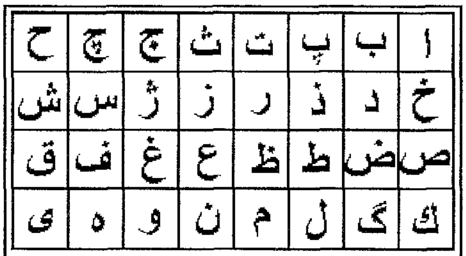

Fig. 1 : Farsi Character Set

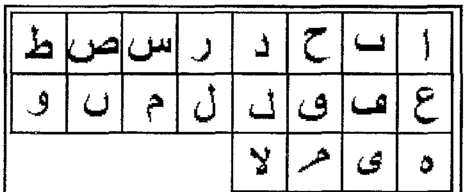

Fig. 2 : Main Body Of Farsi Characters

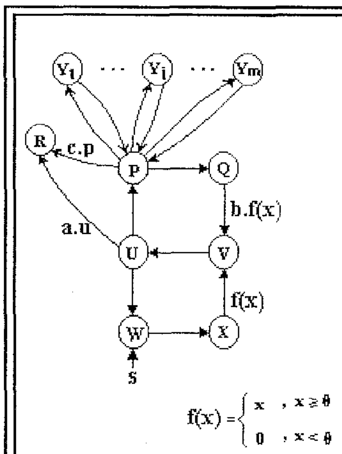

Fig. 3 : Basic ART2 Architecture

Legendre Moment ■ Zernike Moment ● Pseudo Zernike Moment ◆

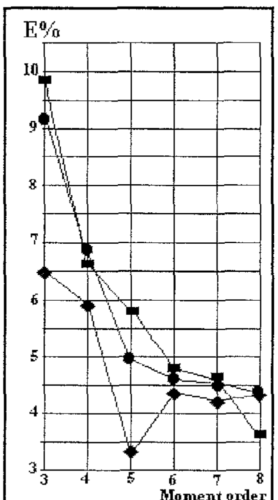

Fig. 4 : Error Rate In Trianing Phase

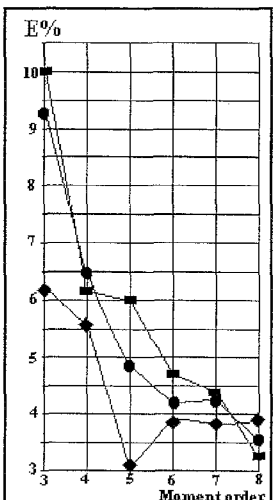

Fig. 5 : Error Rate In Testing Phase

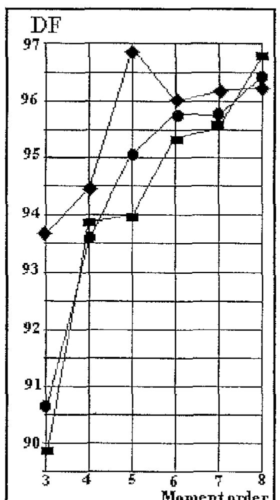

Fig. 6 : Discrimination Factor In Testing Phase