Krawtchouk Moment Feature Extraction for Neural Arabic Handwritten Words Recognition

Anass El affar† Khalid Ferdous†† Abdeljabbar Cherkaoui††† Hakim El fadili†††† and Hassan Qjidaa†††††

University of science dar el mehrazi Fes, department de physique lessi, BP 1796, Fes morocco

Summary
This paper proposes a new approach investigating the application of moment method to evaluate a set of candidate features and to select an informative subset to be used as input data for a neural network classifier. The first step (pre-processing) of proposed method takes into account the discriminative properties of invariant krawtchouk moments. The second step (recognition) is achieved by using multilayer feedforward neural network (MFNN) as a classifier with the stochastic back propagation as a learning algorithm. Finite vectors obtained as a result in the pre-processing phase are then fed into the neural network system. We demonstrate experimentally that the choice of a krawtchouk moment subset which contains sufficient and discriminative information about the input pattern is crucial in the convergence of the neural network training algorithm to a satisfactory performance level. The proposed method has been tested on the well known IFN/ENIT database of Arabic handwritten words. It produces excellent and encouraging result by reducing the computational burden of the recognition system and presenting a high recognition rate with good generalization ability.

Key words:
Method of moments, invariant krawtchouk moments, multilayer feedforward neural network, Arabic handwritten recognition

1. Introduction
Artificial neural networks, and especially multilayer perceptrons (MLP), have shown good capabilities in performing handwritten character recognition. However, their performance is strongly affected by the quality of the representation of the characters. This may require a large number of parameters to represent the character, which then results in difficulty in establishing the rules for recognition. In other words the MLPs become difficult to train. Moreover, the greater the size of network, the greater is the computation time. This can greatly restrict the practical use. So, it is necessary to perform efficient features extraction on the one hand, and to optimize the lay-out of the artificial neural network on the other hand.

In fact, the choice of features to represent the patterns is of capital importance due to the fact that they affect several aspects of the pattern recognition problem such as accuracy, required learning time and necessary number of samples [1].

Different features have been used in the context of character recognition, of particular note, the Statistics-based approaches are very important for their use of global information in an image for extracting features [2]. Especially orthogonal moments have been extensively employed for their shift, rotation, and scale invariance and high robustness in the presence of noise, in classification, recognition, target identification and scene analysis [2-5]. In this paper, we focus on the discriminative power of Krawtchouk moments as a global features to characterize patterns and we then propose a new approach which extract: (a) structural moments i.e. moments that can discriminate clearly the original object in the decision space, collecting the maximum of information needed for representing and reconstructing this object, (b) a reduced number of those moments in order to minimize the computation time and the computational complexity of the classifier, because the moment vector obtained determine the input size of the classifier (a MFNN in our case). If the vector size is reduced and predetermined and if moments extracted are greatly discriminative, the classifier performs well the task of decision.

The proposed contribution for object recognition has two steps: preprocessing and recognition. In the first one, we propose a novel method that extracts optimal object features. For this, we introduce the Maximum Entropy Principle (MEP) as a selection criterion [6]. Our objective is to reduce the input dimensionality of the classification problem by eliminating features with low information content or high redundancy with respect to other features. The second step (recognition) is achieved by using multilayer feedforward neural network as a classifier with the stochastic backpropagation algorithm, where finite vectors obtained in the preprocessing phase are used as inputs to it. The method is tested using the well known IFN/ENIT database of handwritten words [19].

In this work, a class of Krawtchouk moments is examined. Nevertheless, the presented results can be extended to other types of orthogonal moments [8], [9]. Our paper is organized as follows: in Section 2, some basic definitions are given to build-up necessary mathematical background, including Krawtchouk moments and their properties. Section 3 points out the discrimination power of Krawtchouk moment. Sections 4
and 5 present the optimal Moment selection method and our MFNN design. Finally, section 6 and 7 deal with the summary of important results and conclusions of the paper.

2. Krawtchouk moment

Krawtchouk moments are a set of moments formed by using krawtchouk polynomials as the basis function set. Krawtchouk polynomials, introduced by Mikhail krawtchouk [20],[22], are a set of polynomials associated with the binomial distribution. In this section, the definitions of krawtchouk and weighted krawtchouk polynomials are first provided followed by krawtchouk moments and krawtchouk moments invariants.

2.1 krawtchouk polynomials

The definition of the n-th order classical krawtchouk polynomial [21] is defined as:

\[ k_n(x, p; N) = \sum_{i=0}^{N} a_{i} \cdot x^i = \sum_{i=0}^{N} a_{i} \cdot \binom{i}{N} \cdot x^i \]

Where \( x, n = 0, 1, 2, ..., N, N > 0, p \in (0,1) \). \( z_{F_i} \) is the hyper geometric function, defined as

\[ z_{F_i}(a, b; c; z) = \sum_{k=0}^{\infty} \frac{(a)_k \cdot (b)_k \cdot z^k}{(c)_k \cdot k!} \]

And \( (a)_k \) is the pochhammer symbol given by

\[ (a)_k = a(a+1)(a+2) \ldots (a+k-1) = \frac{\Gamma(a+k)}{\Gamma(a)} \]

The set of \( (N+1) \) krawtchouk polynomials \( \{k_n(x, p; N)\} \) forms a complete set of discrete basis functions with weight function

\[ w(x; p, N) = \binom{N}{x} \cdot p^x \cdot (1-p)^{N-x} \]

And satisfies the orthogonality condition

\[ \sum_{x=0}^{N} w(x; p, N) \cdot \overline{k}_n(x, p; N) \cdot k_m(x, p; N) = \delta_{nm} \]

Where \( n, m = 1, 2, ..., N \) and

\[ \rho(n; p, N) = (-1)^n \left( \frac{1-p}{p} \right)^n \cdot \frac{n!}{(N)_n} \]

Examples of krawtchouk polynomials up to the second order are

\[ k_0(x, p, N) = 1 \]
\[ k_1(x, p, N) = 1 - \left( \frac{1}{N-p} \right) \cdot x \]
\[ k_2(x, p, N) = 1 - \left( \frac{2}{N-p} \right) \cdot \frac{1}{N(N-1)p^2} \cdot x + \left( \frac{1}{N(N-1)p^2} \right) \cdot x^2 \]

2.2 weighted krawtchouk polynomials

The conventional method of avoiding numerical fluctuations for moment computations is by means of normalization by the norm. The normalized krawtchouk polynomials with respect to the norm \( K_n(x; p, N) \) is defined as:

\[ \xi_n(x; p, N) = \frac{k_n(x; p, N)}{\sqrt{\rho(n; p, N)}} \]

The set of weighted Krawtchouk polynomials \( \{\xi_n(x; p, N)\} \) is defined by:

\[ \xi_n(x; p, N) = k_n(x; p, N) \cdot \sqrt{\frac{w(x; p, N)}{\rho(x; p, N)}} \]

Such that the orthogonality condition becomes

\[ \sum_{x=0}^{N} \xi_n(x; p, N) \cdot \overline{\xi}_m(x; p, N) = \delta_{nm} \]

2.3 krawtchouk moments

Krawtchouk moments have the interesting property of being able to extract local features of an image. The krawtchouk moments of order \( (n+m) \) in terms of weighted krawtchouk polynomials, for an image with intensity function, \( f(x, y) \), is defined as

\[ Q_{nm} = \sum_{x=0}^{N} \sum_{y=0}^{M} \xi_n(x; p, N-1) \cdot \xi_m(y; p, M-1) \cdot f(x, y) \]

The parameters N and M are substituted with N-1 and M-1 respectively to match the NxM pixel points of an image. The krawtchouk moment corresponding to \( n = m = 0 \) is the weighted mass of the image, i.e.,

\[ Q_{00} = \sum_{x=0}^{N} \sum_{y=0}^{M} \sqrt{w(x; p, N-1) \cdot w(y; p, M-1) \cdot f(x, y)} \]

By solving (8) and (9) for \( f(x, y) \), the image intensity function can be written completely in terms of the krawtchouk moments, i.e.,
One way of interpreting the above equation is that the image intensity function can be represented as a series of weighted krawtchouk polynomials weighted by the krawtchouk moments. If the moments are limited to order $p < 2N-2$, the series is truncated to

$$f(x, y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \phi(n-m, m, x, y)$$

(12)

Where if $S_N = \{0, 1, 2, ..., N-1\}$ (see (13) on the right of the page). Observe from (9) that krawtchouk moments are in fact the inner product of $f(x, y)$ and $k_n(x; p_1, N-1)k_m(y; p_2, N-1)$. Therefore, the appropriate selection of $p_1$ and $p_2$ enables local features at different positions of the image to be extracted by the lower order krawtchouk moments. Using Parseval’s theorem it can be shown that

$$\sum_{n=0}^{N-1} \sum_{m=0}^{M-1} [f(x, y)]^2 = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} [Q_{nm}]^2$$

(14)

As the lower order krawtchouk moments store information of a specific region of interest of an image, the higher order moments store information of the rest of the image. Therefore, by reconstructing the image from the lower order moments and discarding the higher order moments, a subimage can be extracted from the subject image. It is also evident that for each additional moment used in reconstructed image is reduced, that is

$$\Delta E = \Delta \left[ \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} [f(x, y) - \bar{f}(x, y)]^2 \right] = -[Q_{nm}]^2$$

(15)

Where $E$ is the square error and $Q_{nm}$ the additional moment. It follows that, if all $N \times M$ moments are used, the image is perfectly reconstructed. For the rest of the paper, we assume the case of $N=M$.

2.4 krawtchouk moment invariant

If the geometric moments of an image with image intensity function $f(x, y)$ is defined using discrete sum approximation as

$$M_{nm} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} x^n y^m f(x, y)$$

(16)

Then the standard set of geometric moment invariants, which are independent to rotation, scaling and translation [1] can be written as

$$\nu_{nm} = M_{00}^{-1} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \left[ (x - \bar{x}) \cos \theta + (y - \bar{y}) \sin \theta \right]^n \times \left[ (y - \bar{y}) \cos \theta - (x - \bar{x}) \sin \theta \right]^m f(x, y)$$

(17)

Where

$$\gamma = \frac{n + m}{2} + 1$$

(18)

$$\bar{x} = \frac{M_m}{M_00}$$

(19)

$$\bar{y} = \frac{M_{01}}{M_00}$$

(20)

$$\theta = \frac{1}{2} \tan^{-1}\frac{2\mu_{10}}{\mu_{02} - \mu_{20}}$$

(21)

And $\mu_{nm}$ are the central moments defined in [1] as

$$\mu_{nm} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^n (y - \bar{y})^m f(x, y) dx dy$$

(22)

$$\phi(k, l, x, y) = \left\{ \begin{array}{ll}
Q_{nm}k_n(x; p_1, N-1)k_m(y; p_2, M-1) & k \in S_N, l \in S_M \\
0 & \text{otherwise}
\end{array} \right.$$ (13)

The value of $\theta$ according to (21) is limited to $-45^\circ \leq \theta \leq 45^\circ$.

The krawtchouk moments of $\bar{f}(x, y) = \bar{w}(x, y) \nu_{ij} \bar{f}(x, y)$ can be written in terms of geometric moment as

$$Q_{nm} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} k_n(x)k_m(y)\bar{f}(x, y)$$

$$= [\rho(n)\rho(m)]^{1/2} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} k_n(x)k_m(y)f(x, y)$$

$$= [\rho(n)\rho(m)]^{1/2} \sum_{i=0}^{n} \sum_{j=0}^{m} a_{i, n, p} a_{j, m, p} M_{ij}$$

(23)

Where $\{a_{k, n, p}\}$ are coefficients determined by (1). Hence, $Q_{nm}$ is a linear combination of geometric moments, $M_{ij}$ up to order $i = n$ and $j = m$, weighted by coefficients $\{a_{k, n, p}\}$. Notice that (23) transforms the no orthogonal geometric moments to form the orthogonal krawtchouk moments. Notice that the normalized image according to (17) does not fall inside the domain of $[0, N-1] \times [0, N-1]$, as required by krawtchouk moments; therefore, it is modified to
\[ \tilde{v}_{\nu} = \sum_{p,q} \left( n \left( m \left( \frac{N^2}{M_{\infty}} \right) \right) \right) \times \left( \frac{N}{2} \right)^{p+q} \nu_{pq}, \tag{25} \]

The centroid of the image is now shifted to \((N/2, N/2)\), and the image is scale-normalized such that \(\tilde{v}_{\nu} = (N^2/2)\).

Since \(\tilde{v}_{\nu}\) is a linear combination of \(\nu_{pq}\), \(\nu_{mn}\) for odd \(n\) and/or \(m\) does not vanish for symmetrical images; hence, (24) solves the symmetrical problem addressed by Palaniappan et al. in [23]. The new set of moments can be formed by replacing the regular geometric moments \(\{M_{\infty}\}\) by their invariant counterparts \(\{\nu_{mn}\}\). From (23) we have

\[ \tilde{Q} = \rho(n)\rho(m) \sum_{j=0}^{n} \sum_{j=0}^{m} a_{j,n} a_{j,m} \tilde{v}_{ij} \tag{26} \]

Note that the new set of moments is rotation, scale and translation invariant. We shall designate this set of moments as krawtchouk moment invariants in the rest of this paper. Some examples of them are

\[ \tilde{Q}_{\nu} = \Omega_{\nu} \tilde{v}_{\nu}, \]

\[ \tilde{Q}_{ij} = \Omega_{ij} \left[ \tilde{v}_{ij} - \frac{1}{(N-1)p_i} \tilde{v}_{0j} \right] \]

\[ \tilde{Q}_{ij} = \Omega_{ij} \left[ \tilde{v}_{ij} - \frac{1}{(N-1)p_j} \tilde{v}_{0i} \right] \]

\[ \tilde{Q}_{ij} = \Omega_{ij} \left[ \tilde{v}_{ij} - \frac{1}{(N-1)p_i} \tilde{v}_{0j} + \frac{1}{(N-1)p_j} \tilde{v}_{0i} \right] \]

Where \(\Omega_{ij} = [\rho(n; p_1, N-1)p(m; p_2, N-1)]^{-1}\). Note that, in our case, we set the parameters to \(p_1 = p_2 = 0.5\), so that the emphasis of the moments will be at the center of the image. This is consistent with the fact that (24) normalizes the image and shift the centroid to the center of the \([0, N-1] \times [0, N-1]\) plane.

3. Discrimination power of krawtchouk moments

In this section we focus on the discrimination power of Krawtchouk moments, for this let’s consider a subset of 1000 images representing the Arabic hand-written from the IFNENIT database. The projection of those words onto the Krawtchouk polynomials is investigated in order to provide an exploratory data analysis.

Figure 1 represents the projection of the Krawtchouk moment of two patterns representing the Arabic hand-written words: "هاري" and "نا صور" on the 2-dimensional space formed by \((Q_{00}, Q_{20})\) (figure 1(a)), and by \((Q_{00}, Q_{20})\) (figure 1(b)) respectively.

We can easily see, from the presented figures, that this moments sub-space clearly categories the input data into two classes representing each of the digits under study. Those simulations show that the investigated Krawtchouk moment highly discriminate the previous digits in the decision space. The next section deals with the selection of the subset of moment to represent the patterns under investigation in the input database.

4. Neural network design

Neural network is widely used as a classifier in many handwritten character recognition systems [13], [14]. Also, due to the simplicity, generality, and good learning ability of neural networks, these types of classifiers are found to be more efficient [15]. In this paper, multilayer feed forward neural network (MFNN) is used to classify the patterns. In our algorithm, the stochastic gradient algorithm as a minimization procedure is used during the learning phase. The weights are updated on the basis of a single sample. With this procedure the parameter vector
fluctuates around an average trajectory, but usually converges considerably faster than regular gradient descent and second order methods on large training sets with redundant samples [7], [16]. The input of the MFNN are feature vectors derived from the proposed feature extraction method described in the previous section. The number of nodes in the output layer is set to the number of Arabic words classes. In order to avoid stagnation of the gradient descent method in local minima, we use the momentum technique in the weight update procedure. Experiments were conducted using the initial weight vectors that have been randomly chosen from a uniform distribution in (-1, 1), this weight range has been used in [17], [18].

5. Experimental results

In order to illustrate our approach, the method is tested using the IFNENIT database of Arabic handwritten words [19], which has a training set of 26,459 city words with 300 dpi binary handwritten words (town/village names). Figure 2(a) shows some examples randomly picked from the training set of IFNENIT database words.

![Figure 2(a)](image)

Table 1 shows optimal orders obtained by our moment extraction algorithm. In this study, the classifier error rate \( \tau \) (%) is considered as the number of misclassifications in the training (test) phase over the total number of training (test) images.

Table 2 Different test error rates obtained with different MFNN architectures, where the feature subset size is 66 corresponding to order 10.

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Test Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>66,60,10</td>
<td>7.8</td>
</tr>
<tr>
<td>66,90,10</td>
<td>5.62</td>
</tr>
<tr>
<td>66,100,10</td>
<td>5.32</td>
</tr>
<tr>
<td>66,120,10</td>
<td>4.75</td>
</tr>
<tr>
<td>66,140,10</td>
<td>4.21</td>
</tr>
</tbody>
</table>

Table 3 Test error rates of some methods reported in the literature [17]. (HU: hidden Units)

<table>
<thead>
<tr>
<th>Methods</th>
<th>Test Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pairwise linear classifier</td>
<td>8.5</td>
</tr>
<tr>
<td>2-layer NN, 1000 HU</td>
<td>5.3</td>
</tr>
<tr>
<td>2-layer NN, 1000 HU, [distortions]</td>
<td>4.6</td>
</tr>
<tr>
<td>2-layer NN, 300 HU, [distortions]</td>
<td>4.2</td>
</tr>
<tr>
<td>3-layer NN, 300+100 HU</td>
<td>4.1</td>
</tr>
<tr>
<td>3-layer NN, 500+150 HU</td>
<td>3.5</td>
</tr>
</tbody>
</table>
6. Conclusion

In this paper an efficient feature extraction technique is developed, based on the orthogonal moments using Invariant Krawtchouk moments. We have focused on the discrimination power of Invariants Krawtchouk moments. We have shown that the proposed Krawtchouk moment extraction method with MFNN classifier reduce the computational burden of the recognition system in terms of the total number of layers and nodes, while showing an improved performance in terms of recognition rate and generalization ability. Experimental results show that the recognition rate with only one layer and 140 nodes reaches 3.21% on the test set of IFNENIT database.

References

Anass El Affar completed his master ESSID from Dept of physics, Sidi Mohammed Ben Abdellah university, Fès, Morocco in 2003. He successfully completed his undergraduate project on image processing under the supervision of Pr.H.Qjidaa.

Khalid Firdous completed his master ESSID from Dept of physics, Sidi Mohammed Ben Abdellah university, Fès, Morocco in 2003. He successfully completed his undergraduate project on image processing under the supervision of Pr.H.Qjidaa.

Abdeljabar Cherkaoui completed his master ESSID from Dept of physics, Sidi Mohammed Ben Abdellah university, Fès, Morocco in 2003. He successfully completed his undergraduate project on image processing under the supervision of Pr.H.Qjidaa.

Hakim El Fadili completed his master ESSID from Dept of physics, Sidi Mohammed Ben Abdellah university, Fès, Morocco in 2003. He successfully completed his undergraduate project on image processing under the supervision of Pr.H.Qjidaa.

Hassan Qjidaa received his M.Sc and PhD in applied physics from Claude Bernard University of Lyon France in 1983 and 1987 respectively. He got the Pr. Degree in Electrical Engineering from Sidi Mohammed Ben Abdellah University, Fès, Morocco 1999. He is now an associate professor in the Dept. of Physics in Sidi Mohammed Ben Abdellah University, Fès, Morocco. His main research interest include Image Manuscripts Recognition, cognitive Science, Image Processing, Computer Graphics, Pattern Recognition, Neural Networks, Human-machine Interface, Artificial Intelligence, Robotics and so on.