Pattern Recognition with Rotation Invariant Multiresolution Features

S. Rodtook, and S.S. Makhanov

Information Technology Program, Sirindhorn International Institute of Technology
Thammasat University, Rangsit Campus, Pathumthani 12121, Thailand
(E-mail: sittisak@siit.tu.ac.th, makhanov@siit.tu.ac.th)

Abstract: We propose new rotation moment invariants based on multiresolution filter bank techniques. The multiresolution pyramid motivates our simple but efficient feature selection procedure based on the fuzzy C-mean clustering, combined with the Mahalanobis distance. The procedure verifies an impact of random noise as well as an interesting and less known impact of noise due to spatial transformations. The recognition accuracy of the proposed techniques has been tested with the preceding moment invariants as well as with some wavelet based schemes. The numerical experiments, with more than 30,000 images, demonstrate a tangible accuracy increase of about 3% for low noise, 8% for the average noise and 15% for high level noise.

Keywords: Pattern recognition, rotation moment invariant, wavelets, filter bank scheme, feature selection, fuzzy C-mean, Mahalanobis distance.

1. INTRODUCTION

The first region based moment invariants introduced by Hu [1],[2] are projections of the image onto the monomial functions. The moments are believed to be reliable for complex shapes because they involve not solely the contour pixels but all the pixels constituting the object. However, a dramatic increase in complexity associated with the relevant projections makes Hu's moments impractical. Besides, the redundancy of Hu's moments noticed in [2] has clearly indicated a need for further research. Shortly after Hu's paper, a variety of invariant moments has been proposed and analyzed [2]-[8]. The major developments are characterized by the Legendre moments [1],[2], the Zernike moments [2],[3], the Fourier-Mellin moments [2], the Complex moments [2],[4] and the Tchebichef moments [5]. Finally, Shen [6] introduced a rotationally invariant moment representing the image by projections onto wavelets. It has been demonstrated that such wavelet invariants may ensure a higher classification rate.

Multiresolution pyramid is a very well-known procedure. However, to the best of our knowledge, it has not been applied to construct rotation moment invariants. Therefore, we propose to develop the idea of the wavelet-based moments by introducing the filter bank representation and by analyzing its performance. Since the Mallat-like expansion is always overcomplete, the features are selected by the Fuzzy C-mean (FCM) clustering endowed with the Mahalanobis distance and elimination of redundant and noise sensitive features. The objects are represented by the FCM clusters. A minimum of the FCM cost function corresponds to a better discriminative set. Even in the absence of noise induced by physical devices there always exists a noise due to finite resolution of the image subjected to the spatial transformations [7]. The moment invariants should be evaluated by the response not only to random high frequency noise but also to low frequency noise of the rotations and scaling [7].

Recognition rate of the algorithm has been tested by 30,000 different images and compared with the Zernike moments, the Fourier-Mellin moments as well as with a wavelet based representation proposed by Shen [6]. Our proposed techniques provide a significant accuracy increase ranging from 3% to 15%.

2. ROTATIONALLY INVARIANT MOMENTS

A general moment M of a function $f(r,\theta)$ with respect to a moment function $F(r,\theta)$ in the polar coordinate system with the origin at the centroid of the object is defined

$$M = \int_0^{2\pi}\int_0^1 f(r,\theta)F(r,\theta)rdrd\theta$$

We assume that $F(r,\theta) = R(r)G(\theta)$, where $R(r)$ denotes a basis function such as the Zernike polynomial, and $G(\theta)$ an angular function. Taking $G(\theta) \equiv G_\Gamma(\theta) = e^{i\Gamma\theta}$ for some Γ provides the rotational invariance. Note that if Γ is considered as a continuous variable the integral with regard to θ is nothing else that the circular Fourier transforms. Usually (but not necessarily), in the theory of rotational invariant moments, Γ is an integer [6] called the angular order. We represent the above integral by

$$M_\Gamma = \int_0^1 R(r)S_\Gamma(r)rdr$$

where $S_\Gamma(r) = \int_0^{2\pi} f(r,\theta)G_\Gamma(\theta)d\theta$.

Note that if $\tilde{M}_\Gamma$ is a moment of the rotated image $f(r,\theta+\phi)$, where ϕ is the angle of rotation, then $\tilde{M}_\Gamma = e^{i\Gamma\phi}M_\Gamma$. Therefore, $|\tilde{M}_\Gamma| = |M_\Gamma|$. Thus, rotation of the object affects the phase but not the magnitude. Furthermore, from the viewpoint of functional analysis, each object is represented by an infinite and unique set of the invariants if functions $R(r)$ constitute a basis in the appropriate functional space.

A wavelet basis has a number of advantages since it could be adapted to the spectrum as well as to the spatial properties of a particular set of objects. In [6] the set of the radial functions is given by $R(r) = \psi_{m,n}(r) = (1/\sqrt{m}) * \psi(r-n)/m$, where $\psi(r)$ is the mother wavelet, *m* the dilation parameter (the scale index) and *n* the shifting parameter.

We call projections of $S_\Gamma(r)r$ onto $R_{m,n}(r)$ the wavelet moments and denote them by $M_{m,n,\Gamma}$. From the point of view of multiresolution analysis, such projections correspond to the "details" associated with the high frequency part of the object shape which are usually sensitive to noise.

3. FILTER BANK MOMENT INVARIANTS

In this section we introduce new moment invariants based on the filter bank. In the case of discrete orthogonal wavelets the low-resolution coefficients may be calculated from higher resolution coefficients by a scheme called the filter blank. The QMF is a fast algorithm first proposed by Mallat [8] and extended to the biorthogonal by Unser et al [9]. The approximation and the detail wavelet moments are constructed as follows

$$A_{m,n,\Gamma} = \sum_q H_{2n-q} A_{m+1,q,\Gamma}$$

$$D_{m,n,\Gamma} = \sum_q G_{2n-q} A_{m+1,q,\Gamma}$$

$m = V_0 - 1, V_0 - 2, ..., 0$, where V_0-1 is the finest resolution level. H, G are the so-called finite impulse response filters [8], $A_{V_0,n,\Gamma} = S_\Gamma(r_n) r_n$, $r_n = n\frac{1}{K}$, $n = 1, 2, ..., K$. It is not hard to demonstrate that $|A_{m,n,\Gamma}|$ and $|D_{m,n,\Gamma}|$ are rotation invariants for any Γ.

4. SELECTING THE FILTER BANK INVARIANTS

Selection of features [10] is a crucial step for any shape recognition system. The mutliresolution moment invariants imply that for dissimilar objects the features should be taken mostly from the approximation coefficients; however, for similar objects one should employ the details. In order to find the best combination of the approximation and the detail coefficients we first, we examine the features individually and discard those with a low discriminatory capability. Further selection is done by analyzing combinations of the features. The entire feature selection procedure is given below.

1. Discard the noise-sensitive angular orders by considering the least square error type given by

$$error(\Gamma) = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\left|S_\Gamma(r_k)^{i,Template}\right| - \left|S_\Gamma(r_k)^{i,j}\right|\right)^2}{I\,J\,K}$$

where I is the number of classes, J the number of objects in each class and $S_\Gamma(r_k)^{i,Template}$ the circular Fourier transform of the template associated with class i. The resulting set $\Gamma^* = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_L\}$ is fed to the next step of the procedure.

2. For each implement the QMF scheme as illustrated below

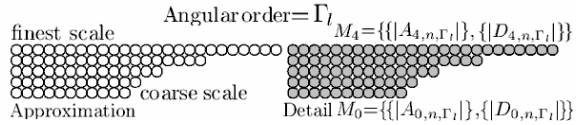


Fig. 1 Pyramidal filter bank moment invariants.

3. Reduce the dimension of the feature space by analyzing the features individually. We threshold the features using a statistical testing ANOVA [11]. We use a one-way ANOVA with a randomized complete block design to verify the assumption $\mu_1 \neq \mu_2 \neq \ldots \mu_i \neq \ldots \neq \mu_I$, where μ_i is the mean-feature of class i.

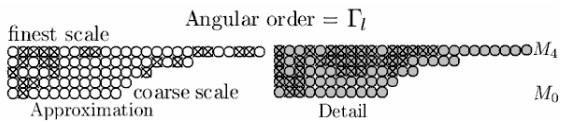


Fig 2 Analysis of the individual filter bank moment invariants.

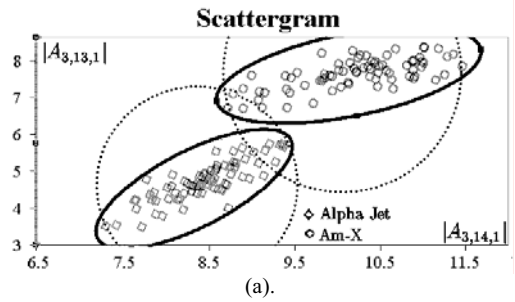
4. Analyze combinations of the features. At this stage the multiresolution analysis is combined with the FCM technique [12]. First of all, the features must be normalized [13] otherwise the FCM cost function could be larger for a better feature set. We consider all possible combinations of the features M_{a,Γ_i} at scale a; then $F_a = \{M_{a,\Gamma_1}, M_{a,\Gamma_2}, ..., M_{a,\Gamma_L}\}$. Next, we consider combinations of the features selected from the scales *a, b, c, d, e…* as follows

$$F_{ab} = \{\{M_a M_b\}_{\Gamma_1}, \{M_a M_b\}_{\Gamma_2}, ..., \{M_a M_b\}_{\Gamma_L}\}$$
$$F_{abc} = \{\{M_a M_b M_c\}_{\Gamma_1}, \{M_a M_b M_c\}_{\Gamma_2}, ..., \{M_a M_b M_c\}_{\Gamma_L}\}$$
$$F_{abcd} = \{\{M_a M_b M_c M_d\}_{\Gamma_1}, ..., \{M_a M_b M_c M_d\}_{\Gamma_L}\}$$
$$F_{abcde} = \{\{M_a M_b M_c M_d M_e\}_{\Gamma_1}, ..., \{M_a M_b M_c M_d M_e\}_{\Gamma_L}\}$$

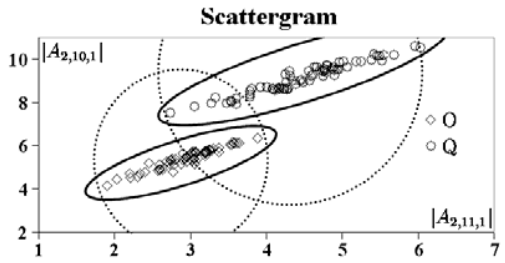
…, etc

The discriminatory capability of a set is evaluated by minimizing the FCM cost function. A minimum of the function corresponds to a better set. As mentioned before, the filter bank moment invariants are redundant and correlated, therefore, the FCM algorithm involves the Mahalanobis distance [14]. It also provides better separability. Consider scattergrams $|A_{3,13,1}|/|A_{3,14,1}|$ associated with two similar aircrafts Alpha Jet and Am-X and $|A_{2,10,1}|/|A_{2,11,1}|$ associated with two upper case English letters "*O*" and "*Q*" (Fig 3 (a) and 3(b) respectively). The training classes form an elliptical shape. The circles representing the classes in terms of the Euclidean distance overlap (see Fig 3). However, the two ellipses representing the Mahalanobis distance are separable. Moreover, even if the sets are separable in the both metric, the Mahalanobis metric usually requires less FCM iterations (see Fig 4 (a) and (b)).

Finally, once an appropriate feature set has been selected, the classification templates are automatically found as the centroids of the FCM clusters.

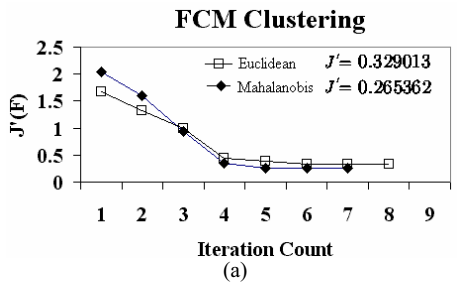


(a).

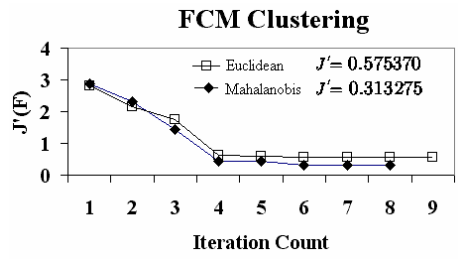


(b).

Fig 3. Scattergrams of Alpha Jet and Am-X and the characters O and Q corrupted by 0-2% impulse nose, (a). Aircrafts. (b). Characters.



(a)



(b).

Fig. 4. Convergence of the FCM cost function, two rotated/scaled objects corrupted by 0-2% impulse noise, (a). Aircrafts: $F = \{|A_{3,13,1}|, |A_{3,14,1}|\}$, (b). Characters *O* and *Q*, $F = \{|A_{2,10,1}|, |A_{2,11,1}|\}$

5. EXPERIMENTAL RESULTS

We evaluate performance of the proposed algorithm by two datasets. The first data consists of 37,500 noisy images [15] based on fifteen basic aircraft silhouettes: Alpha Jet, Am_x, Jaguar, Hawk, An-12 Cub, An-24 Coke, An-32 Cline, C-130 Provider, C-137 Hercules, G-222, MB-326, MB-339A, Mig-29, MiG-17 and Jastreb. Each silhouette produces 1600 training images and 900 testing images. Our second dataset based on online database NIST [16], consists of machine-printed characters, 9000 upper case English letters (Bold, Courier). We use 5800 letters for training and 3200 for testing. The both datasets are degraded by an impulse noise varying from 1 to 8% and the transformation noise. We perform the experiments by the filter blank moment invariants obtained by means of the cubic B-splines. The orthogonal Daubiech wavelets of order 2, 4 and 6 and the Coiflet wavelets were tested as well. Although the orthogonal wavelets easily allow to reconstruct the image, the B-splines are always performing slightly better. We also analyzed the wavelet moment invariants introduced by Shen [6]. As mentioned before, the Shen's invariants obtained by projecting $\psi_{m,n}(r)$ onto $S_\Gamma(r)r$ correspond to $D_{m,n,\Gamma}$.

Denote our proposed algorithm by QMF-FCM-M in the case of the Mahalanobis distance and by QMF-FCM-E in the case of the Euclidean distance. We use the notation "FCM" to indicate our feature selection algorithm and the notation "I" if the features were selected individually by mean of Shen's algorithms (see for instance [6]). For example, "Shen-I-E", means "the Shen's invariants with individual selection in terms of the Euclidean distance".

The comparison of an average classification rate of the proposed QMF-FCM-M versus the most popular moment invariants is shown in Table 1. Table 1 includes degradation by all types of noise, rotation, translation, scaling and random noise. Besides, in the case of NIST we consider an interesting effect of the boundary noise appearing after separation of touching letters by means of dilation. Consider Tables 1. Shen-I-E applied to the NIST symbols has 88.4% average recognition rate, whereas our method provides 94%. The table shows that every component of the algorithm is almost equally important. Namely, combining the QMF with the FCM shows a 1.5% increase whereas the Mahalanobis distance increases the recognition rate further by 2%. Differentiation by the intensity and the type of noise given in Tables 2-5 reveals that our algorithm almost always overperforms Shen's invariants especially when they are based on the individual feature selection and the Euclidean distance. The efficiency of the algorithm with the reference to the preceding techniques becomes apparent when increasing the noise intensity. The most impressive result is an almost 28 % absolute increase (45% relative increase) with regard the Fourier Mellin (FM) invariants in the case of the aircraft silhouettes degraded by 6-8% impulse noise and the rotation noise. Tables 4 and 5 exemplify the experiments with the NIST printed characters. Note that the rotation and segmentation noise display much more significant impact on the characters since the centroids of the characters often lie outside the character body. Consequently, the centroids are much more sensitive to the noise.

Table 1. Average classification rates.

Algorithm	Classification rates	
	Aircraft silhouettes	Upper case characters
QMF-FCM-M	92.8 %	94%
QMF-FCM-E	89.4%	90.9%
Shen- FCM-M	91.2 %	92.6%
Shen-FCM-E	86.9%	89%
Shen -I –E	84.7%	88.4 %
Zernike -I– E	82.8%	86.1 %
FM- I –E	77.1%	80.3

Table 2. Aircraft images, impulse noise.

Impulse noise ratio η%	$0 \le \eta < 2$ %	$2 \le \eta < 4$ %	$4 \le \eta < 6$ %	$6 \le \eta < 8$ %
QMF-FCM-M	98.7	93.9	86.2	67.3
Shen- FCM-M	98.1	92.0	82.7	61.9
Shen- I-E	96.4	86.2	73.3	53.8
Zernike -I– E	95.9	85.1	71.8	51.6
FM- I –E	90.7	78.7	59.5	43.7

Table 3. Aircraft images, impulse noise combined with rotation and scaling.

Impulse noise ratio η%	$0 \le \eta < 2$ %	$2 \le \eta < 4$ %	$4 \le \eta < 6$ %	$6 \le \eta < 8$ %
QMF-FCM-M	96.3	91.6	82.3	62.6
Shen- FCM-M	95.4	90.3	78.4	55.0
Shen- I-E	93.1	82.7	65.6	45.3
Zernike -I– E	91.8	80.9	62.9	41.9
FM- I –E	87.3	71.9	47.1	34.5

Table 4. The NIST characters, impulse noise and segmentation noise.

Impulse noise ratio η%	$0 \le \eta < 1.5$ %	$1.5 \le \eta < 3$ %	$3 \le \eta < 4.5$ %	$4.5 \le \eta < 6$ %
QMF-FCM-M	99.7	94.3	88.3	72.9
Shen- FCM-M	99.2	92.4	85.6	68.8
Shen- I-E	97.7	89.6	80.8	61.4
Zernike -I– E	96.9	88.3	78.4	58.6
FM- I –E	94.8	82.1	70.1	49.8

Table 5. The NIST characters, impulse noise and transformation noise.

Impulse noise ratio η%	$0 \le \eta < 1.5$ %	$1.5 \le \eta < 3$ %	$3 \le \eta < 4.5$ %	$4.5 \le \eta < 6$ %
QMF-FCM-M	97.3	93.6	85.3	67.5
Shen- FCM-M	96.6	91.2	81.9	60.1
Shen- I-E	94.7	86.1	72.8	51.5
Zernike -I– E	92.6	83.9	68.3	46.7
FM- I –E	89.3	77.4	61.2	37.8

6. CONCLUSIONS

The proposed filter bank invariants extend the idea of applying wavelets for rotation invariant pattern recognition. Our approach based on the analysis of the high and the low frequency filter bank coefficients combined with elimination of the redundant features always leads to a tangible improvement of the recognition rate with the reference to the conventional methods. For instance, on average we obtain an increase of about 3% for low noise, 8% for an average noise and 15% for high level noise. A large number of testing images and the variety of the sources of the noise makes it possible to conjecture that the proposed technique performs better than the existing ones for other applications.

REFERENCES

[1] J. Shen, W. Shen and D. Shen, "On Geometric and Orthogonal Moments", *Multispectral Image Processing and Pattern Recognition Series in Machine Perception Artificial Intelligence*, Vol. 44, World Science, pp 17-36, 2001.

[2] C. H. Teh and R. T. Chin, "On Image Analysis by the Methods of Moments", *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 10, No. 4, pp. 496-512, 1988.

[3] S. X. Liao and M. Pawlak, "On the Accuracy of Zernike Moments for Image Analysis", *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 20, No. 12, pp. 1358-1364, 1998.

[4] J. Flusser, "On the inverse problem of rotation moment invariants", *Pattern Recognition*, Vol. 35, pp. 3015-3017, 2002.

[5] R. Mukundan, S.H. Ong and P.A. Lee, "Image analysis by Tchebichef moments", *IEEE Trans. on Image Processing*, Vol. 10, No. 9, pp. 1357–1364, 2001.

[6] D. Shen and H. H. Ip, "Discriminative Wavelet Shape Descriptors for Recognition of 2-D patterns", *Pattern Recognition*, Vol. 32, pp. 151-165, 1999.

[7] S. Rodtook and S.S. Makhanov, "On the Accuracy of Rotation Invariant Wavelet-Based Moments Applied to Recognize Traditional Thai Musical Instruments", *LNAI 2773 Knowledge-Based Intelligent Information and Engineering Systems*, Vol. 1, pp. 408-414, 2003.

[8] S.G. Mallat, "Multi-frequency channel decompositions of images and wavelet models", *IEEE Trans. Image Process*, Vol. 37, No. 12, pp. 2091-2110, 1989.

[9] M. Unser and A. Aldroubi, "Polynomial Splines and Wavelets--a Signal Processing Perspective", *Wavelets-A Tutorial in Theory and Application*, Academic Press, Boston, pp. 91-122, 1992.

[10] A.K. Jain, R.P.W Duin and J. Mao, "Statistical Pattern Recognition: A Review", *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, pp 4-35, 2000.

[11] D. Tipyota, "Analysis of Variance, in: Probability and Statistics", *Cuprint Press*, Bangkok, pp. 329-333, 2002.

[12] M.R Rezaee, P.M Van der Zwet, B.P Lelieveldt, R.J Van der Geest and J.H Reiber, "A Multiresolution Image Segmentation Technique Based on Pyramidal Segmentation and Fuzzy Clustering", *IEEE Trans. Image Processing*, Vol. 9, No. 7, pp. 1238-1248, 2000.

[13] J. Han and M. Kamber, "Data Preprocessing, in: Data Mining Concepts and Techniques", *Morgan Kaufmann Publishers, An Imprint of Academic Press*, London, pp. 114-116.

[14] Deer, P.J.; Eklund, P.W.; Norman, B.D, **"**A Mahalanobis distance fuzzy classifier"**,** *Intelligent Information Systems Conference*, New Zealand and Australia, Vol. 1, pp. 220 – 223, 1996.

[15] J.A. Wickham JR., "Visual Aircraft Recognition, Manual of Headquarters department of the Army", *US Army Air Defense Artillery School*, Washington DC, 1986.

[16] National Institute of Standards and Technology, *Database 8 NIST Machine-Print DB of Gray Scale and Binary Images (MPDB)*, Standard Reference Data Program, http://www.nist.gov/srd/nistsd8.htm.