

# Apriori Data Mining on Rotationally Invariant Multiresolutional Moments for Pattern Recognition

A. Rodtook<sup>1</sup> and S.S. Makhanov<sup>2</sup>

<sup>1</sup>Department of Computer Science, Ramkhamhaeng University, Thailand

<sup>2</sup>Information Technology and Computer Program, Sirindhorn International Institute of Technology, Thammasat University, Thailand

## Abstract

We propose a new feature selection procedure based on a combination of a pruning algorithm, Apriori mining techniques and fuzzy C-mean clustering. The feature selection algorithm is designed to mine on a multiresolution filter bank composed of rotationally invariant moments.

The numerical experiments, with more than 10,000 images, demonstrate an accuracy increase of about 5% for a low noise, 15% for an average noise and 20% for a high-level noise.

**Keywords:** filter bank, the Kullback-Leibler distance, Apriori mining algorithm, fuzzy C-mean clustering.

## 1. Introduction

A popular class of the rotationally invariant features is based on the moment techniques [1]-[6] which are believed to be reliable for complex shapes because they involve not solely the contour pixels as it is the case for the shape descriptors but all the pixels constituting the object.

Furthermore, as opposed to the shape based descriptors, the rotation invariants may not require an accurate segmentation. It is sufficient to only find the occluding circle. After that the procedure processes each pixel irrespectively whether it belongs to the object or not. Besides, practically all the shape descriptors (except the Fourier transform of the boundary curve and its modifications) are not rotationally invariant. The rotations require point matching such as the shape contexts techniques [19], special coordinate systems, etc.

A particular but important class of such moments is based on a spatial-frequency domain representation. First, a circular Fourier transform (the Fourier transform with regard to the angular coordinate) is applied to a circle occluding the object. The result is a complex function of two real variables, the frequency variable and the radius of the circle. Next, the result is sampled with regard to the frequency variable at integer frequencies. The magnitude of every sample is

rotationally invariant. Finally, the samples, which are functions of the radial variable, are represented in terms of an appropriate spatial basis. The coefficients of the Fourier series in this basis constitute the required moments.

The choice of the spatial basis could be critical for pattern recognition. The most popular options are the Zernike polynomials, Mellin monomials (Fourier-Mellin moments) and complex monomials.

Shen and Ip [5] introduced rotationally invariant moments representing the image by projections onto certain wavelet spaces. It has been demonstrated that the wavelet moments may ensure a higher classification rate with the reference to conventional moments applied to discriminate similar objects such as digit 1 and letter ‘1’, etc.

The advantage of the wavelet based moments is the ability to extract features localized in space and frequency. For example, for two similar objects, the magnitude of the difference between the corresponding Zernike moments or Hu’s moments could become smaller than the magnitude of the noise. Consequently, they may overlap and oscillate from sample to sample [5]. However, if the basis function is locally defined in space and frequency it is often possible to extract a local feature which differentiates the two objects. The corresponding basis function should have the support in a neighborhood of the distinctive feature (space localization) and filter the noise frequencies (frequency localization).

However, the wavelet moments designed in [5] do not fully exploit the concept of multiresolution. The wavelet moments are obtained by representing the circular Fourier transform of the object image by the so called detail coefficients which in [5] are considered suitable for recognition of similar objects belonging to different classes. However, even for similar objects the correspondence between the *image details* and the *detail coefficients* is not that straightforward as it may seem. Actually, the image details effect both the detail and the approximation coefficients.

For example, the recognition by detail coefficients may fail when the object is obtained from an original object by adding a single part or making a hole. As opposed to that, our technique uses the entire multiresolution filter bank which allows for more efficient analysis.

The wavelet filter bank is characterized by a large number of coefficients and is always over complete. Therefore, we propose a multi-stage feature selection designed specifically for the multiresolution features to eliminate noise sensitive, redundant and other non important features. First, the algorithm prunes the multiresolution bank using the Kullback-Leibler distance [10], then the wavelet coefficients are analyzed individually by using the standard ANOVA [11]. Finally, combinations of the coefficients are selected using a modified Apriori algorithm [12] combined with the fuzzy C-mean clustering.

We show that the multiresolution analysis combined with our new feature selection algorithm has a better recognition rate as compared with the preceding methods.

## 2. Rotationally Invariant Moments

A general moment of an image  $f(r, \theta)$  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 by

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

Let  $F(r, \theta) = \beta(r) \alpha(\theta)$ , where  $\beta(r)$  denotes a family of radial functions such as the Zernike polynomial, Mellin polynomial etc. The choice of  $\beta(r)$  defines the type of the moment.  $\alpha(\theta)$  denotes an angular function. Taking  $\alpha(\theta) \equiv \omega_q(\theta) = e^{iq\theta}$  provides the rotational invariance. Note that if  $q$  is a continuous variable, then the integral with regard to  $\theta$  is nothing but the circular Fourier transform. In the theory of rotationally invariant moments  $q$  is an integer called the angular order [5]. We present the above 2-D integral by

$$M_q = \int_0^1 \beta(r) \xi_q(r) r dr,$$

where  $\xi_q(r) = \int_0^{2\pi} f(r, \theta) \omega_q(\theta) d\theta$  is a complex function of spatial variable  $r$ . Note that for a fixed  $r$ ,  $\xi_q(r)$  represents the image in the frequency domain whereas for a fixed  $q$  the representation remains spatial.

If  $\tilde{M}_q$  is a moment of the rotated image  $f(r, \theta + \phi)$ , where  $\phi$  is the angle of rotation, then  $\tilde{M}_q = e^{iq\phi} M_q$ . Therefore,  $|\tilde{M}_q| = |M_q|$ . In other words, rotations of the object affect the phase but not the magnitude.

From the viewpoint of functional analysis, each object is represented by an infinite and unique set of moments if the family of functions  $\beta(r)$  constitutes a basis in the appropriate functional space. In the case of wavelets, a wavelet function is given by  $\psi_{m,n}(r) = 2^{m/2} \psi(2^m r - 0.5n)$  [5], where  $m$  the dilation parameter (the scale index) and  $n$  the shifting parameter. The wavelet bases have a number of advantages since they could be adapted to the spectrum as well as to the spatial properties of a particular set of objects. A multiresolution version of the proposed wavelet filter-bank moments will be introduced next.

## 3. Rotationally Invariant Filter Bank Moments

In terms of the multiresolution analysis the sequence of approximating spaces is generated by the so-called scaling functions  $\varphi$  [7],[18] whereas the wavelet functions are employed to represent the orthogonal complements to the approximating spaces called the detail spaces. We define the approximation and the detail moments respectively as follows

$$a_{m,n,q} = \int_0^1 \xi_q(r) r \varphi_{m,n}(r) dr, \quad d_{m,n,q} = \int_0^1 \xi_q(r) r \psi_{m,n}(r) dr.$$

Note that

$$\sum_n a_{m+1,n,q} \varphi_{m+1,n}(r) - \sum_n a_{m,n,q} \varphi_{m,n}(r) = \sum_n d_{m,n,q} \psi_{m,n}(r).$$

The discrete version of the above decomposition was proposed by Mallat [7]. Mallat has shown that the discrete wavelet transform can be performed by using the so-called finite impulse response filters (FIR) to construct a tree structured filter bank. Mallat proposed a quadrature mirror filter (QMF) which corresponds to the orthogonal wavelets. Unser et al [8] extended these techniques to the biorthogonal wavelets. In the framework of the QMF, the approximation and detail filter bank moments are constructed as follow

$$a_{m,n,q} = \sum_k h_{k-2n} a_{m+1,k,q}, \quad d_{m,n,q} = \sum_k g_{k-2n} a_{m+1,k,q},$$

$m = m_0, m_0-1, \dots, 0$  and  $a_{m_0+1,n,q} = \xi_q(r_n) r_n$ ,  $r_n = (n+1)/N$ ,

$h$  and  $g$  are the FIR filters. The output of the filters is down sampled and then a low-frequency output (approximation) is fed to the identical filters shown in Fig. 1.

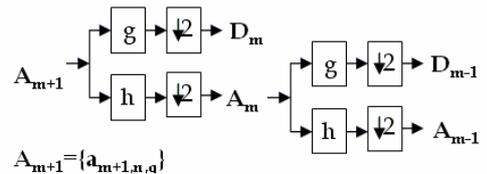


Fig 1: Two-resolution two-band filter bank.

Furthermore, let  $\tilde{f}$  be a rotated image, where  $\phi$  is the angle of rotation, then

$$\begin{aligned}\tilde{a}_{m,n,q} &= \sum_k h_{k-2n} \tilde{a}_{m+1,k,q} = e^{iq\phi} \sum_k h_{k-2n} a_{m+1,k,q}, \\ \tilde{d}_{m,n,q} &= \sum_k g_{k-2n} \tilde{a}_{m+1,k,q} = e^{iq\phi} \sum_k g_{k-2n} a_{m+1,k,q}.\end{aligned}$$

Therefore,  $|a_{m,n,q}|$  and  $|d_{m,n,q}|$  are rotation invariants for any  $q$ .

## 4. Feature Selection

Selection of features is a crucial step for an object recognition system. The aim is to generate the best combination of features that maximizes the recognition rate. We present an algorithm based on examining the filter-bank bands, then the features individually and finally their combinations. A detailed description of the feature selection procedure is given below.

1. Sample the circular Fourier at integer values of  $q$  as follows

$$\xi_q(r) = \int_0^{2\pi} f(r, \theta) e^{iq\theta} d\theta.$$

2. Discard noise-sensitive angular orders by evaluating the least square error type given by

$$E(q) = \frac{\sum_{i=1}^I \sum_{j=1}^J \sum_{n=0}^{N-1} \left( \left| \xi_q(r_n)^{i, Template} \right| - \left| \xi_q(r_n)^{i,j} \right| \right)^2}{IJN},$$

where  $I$  is the number of the classes,  $J$  the number of objects in each class and  $\xi_q(r_n)^{i, Template}$  the circular Fourier transform of the template associated with class  $i$ . This preprocessing allows to discard easily recognizable bad choices. The resulting set  $q'$  is fed to the next step of the procedure.

3. Apply the QMF to  $\xi_q(r_n)r_n$  as illustrated in Fig. 1.

4. Prune the resulting filter-bank. The algorithm finds the best discriminant sub bands which produce well-separated classes. A symmetric version of the Kullback-Leibler distance based on the relative entropy [10] is used to measure the discrimination power of the sub band. A good sub band is the one that minimizes the relative entropy [10] which for the case of two classes  $C_1$  and  $C_2$  is given by

$$\delta_{C_1 C_2}(m, k, q) = 1/2 \sum_n \left( \gamma_{C_1}(m, k, n, q) \log \left( \frac{\gamma_{C_1}(m, k, n, q)}{\gamma_{C_2}(m, k, n, q)} \right) + \gamma_{C_2}(m, k, n, q) \log \left( \frac{\gamma_{C_2}(m, k, n, q)}{\gamma_{C_1}(m, k, n, q)} \right) \right),$$

where

$$\gamma_{C_i}(m, k, n, q) = \sum_{C_i} \left( \left( w_{m,n,q} \cdot w_{m,n,q}^* \right) / \sum_l \left( a_{m+1,l,q} \cdot a_{m+1,l,q}^* \right) \right).$$

$m$  is the scale index,  $k$  the branch of the tree at scale  $m$  ( $k=0,1$ ) and  $n$  the spatial index.  $w_{m,n,q}$  represents the wavelet coefficients as follows

$$w_{m,n,q} = \begin{cases} a_{m,n,q}, & \text{if } k = 0, \\ d_{m,n,q}, & \text{if } k = 1. \end{cases}$$

\* denotes complex conjugate. In the case of  $I$  classes we have  $\delta(m, k, q) = \sum_{i=1}^{I-1} \sum_{j=i+1}^I \delta_{C_i C_j}(m, k, q)$ .

The algorithm evaluates the discrimination power of the sub bands by comparing the Kullback-Leibler distance before and after the split for  $m=0,1, \dots, m_0$  as shown in Fig. 2.

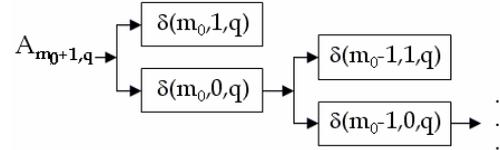


Fig 2: Example of the pruning algorithm for case: if  $\delta(m_0,0,q) > \delta(m_0-1,0,q) + \delta(m_0-1,1,q)$ : keep sub bands  $\delta(m_0-1,0,q)$  and  $\delta(m_0-1,1,q)$ ; otherwise remove both sub bands.

5. Reduce the dimension of the feature space by analyzing the features individually using a statistical testing ANOVA [11]. We use a one-way ANOVA with a randomized complete block design to verify the assumption  $\mu_1 \neq \mu_2 \neq \dots \neq \mu_i \neq \dots \neq \mu_l$ , where  $\mu_i$  is the mean-feature of the class  $i$ . This preprocessing procedure is illustrated in Fig. 3 below

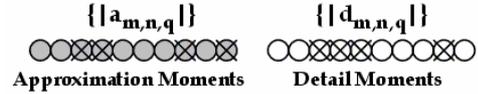


Fig 3: Discarding noise sensitive features in the specified sub bands.

6. Analyze combinations of features. At this stage the result of the multiresolution analysis is fed to the Apriori algorithm. The Apriori selection techniques initially developed for data mining applications reduce the number of combinations appearing when mining for frequent itemsets in large databases. Since the filter bank may produce a large number of features the Apriori algorithm is beneficial in our case as well.

An initial set of the "frequent" features  $L_1$  is selected by ANOVA. It is then used to find  $L_2$  which consists of the best pairs of features taken from  $L_1$ . The set of the discriminant 2-itemsets is then used to find  $L_3$ , and so on.

A set of the candidate  $k$ -itemsets is generated by joining  $L_{k-1}$  with itself  $C_k = L_{k-1} \otimes L_{k-1}$  [12]. The choice of a good combination of the features is based on evaluating a cost function used by analogy with the measure of confidence in the conventional Apriori algorithm [12]. Note that the conventional Apriori

algorithm requires that combinations of features obey the so-called anti-monotonic property, that is, if a set can not pass a test then all of its supersets fail the test as well. We propose the so-called  $\varepsilon$ -anti-monotonic property which makes it possible to pass a local minimum and find a better combination of the features. The confidence in a combination of two features  $A$  and  $B$  with regard to a cost function  $f_C \geq 0$  is evaluated as follows

$$f_C(A \cup B) \leq f_C(A) + \varepsilon \text{ and } f_C(A \cup B) \leq f_C(B) + \varepsilon,$$

where  $\varepsilon$  defines the allowable interval of confidence.

The confidence in a combination of  $n$  features is evaluated using the same principle. The features must be normalized and standardized prior to the selection. Our cost function given by

$$f_C(X) = J(X) \log(N_{\text{miss}} + \mathcal{G}),$$

where  $X$  is a combination of features,  $J$  is the fuzzy C-mean (FCM) cost function [15],[16],[17],  $N_{\text{miss}}$  is the number of training patterns that have been incorrectly clustered,  $\mathcal{G} > 1$  a prescribed constant to eliminate the singularity  $\log(0)$ .

Finally, the best combination is the  $k$ -itemset which minimizes the cost function above.

The Apriori technique applies as follows

- find the best feature set from each particular resolution,
- find the best feature set from the entire multiresolution analysis,
- find the best feature set from the entire set of angular orders  $q'$ .

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

Note that the clustering can be executed in either unsupervised [16] or supervised mode [17] (see the next section).

## 5. Experimental Results

We evaluate the performance of the proposed algorithm by two datasets. The first dataset consists of 14,000 gray level images based on ten Thai musical instruments [6],[13] (see Fig.4). Each instrument produces 950 training images and 450 testing images. The second dataset based on an online database NIST [14], consists of machine-printed characters, namely, 11000 upper case English letters (Bold, Courier). We use 7000 letters for training and 4000 for testing. All of the datasets are degraded by an impulse noise varying from 0% to 8%, a transformation noise (such as rotation and scaling noise) [6] and a boundary noise. The rotation noise affects the NIST characters and the Thai musical instruments significantly since the centroid of the characters and the instruments often

lie outside the object body. Consequently, the centroids are much more sensitive to the noise.

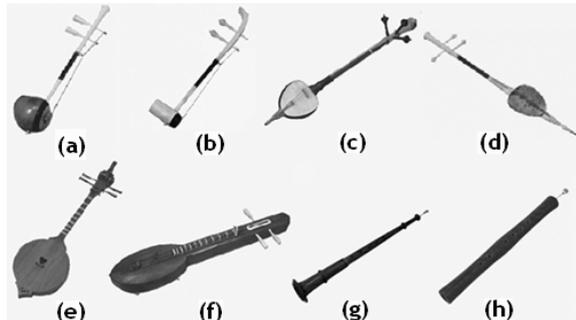


Fig. 4: Gray level images of Thai musical instruments (a). "SAW DUANG" (Fiddle), (b). "SAW OU" (Fiddle), (c). "SALOR" (Lute), (d). "SAW SAM SAI" (Fiddle), (e). "SUENG" (Lute), (f). "JAKAE" (Lute), (g). "PEE CHAWA" (Pipe), (h). "PEE NOKE" (Pipe), (i). "RANAD TUM" (gamelan), (j). "RANAD AKE" (gamelan).

We will discuss experiments with the B-spline wavelets, however, the orthogonal wavelets such as the Daubechies wavelets 2, 4 and 6 and the Coiflet wavelets [18] were tested as well. The biorthogonal B-splines with underlying symmetric FIR filters were always performing slightly better.

Denote our proposed algorithm by QMF-P-AA-FCM-S in the case of FCM with partial supervision [17], by QMF-P-AA-FCM-M in the case of unsupervised FCM endowed with the Mahalanobis distance [16] and by QMF-P-AA-FCM-E in the case of the Euclidean distance. "P-AA" denotes a combination of the pruning algorithm, the standard ANOVA and the modified Apriori mining, "IV" corresponds to individual selection based on the between-to within-class variance ratio [5].

The comparisons of an average classification rate of the proposed QMF-P-AA-FCM-S versus the most popular moment invariants are shown in Table 1. Table 1 includes degradation by all types of noise: rotation, translation, scaling and a random impulse noise,  $\eta$  denotes the noise intensity in %.

Table 1 shows advantages of our approach. For instance, "Shen-IV" [5] applied to the musical instruments has 81.01% average recognition rate, whereas our method provides recognition rate of about 93.24%. The table shows that every component of the algorithm is almost equally important. Namely, combining the QMF with the FCM-E shows a 3% increase. Adding the Mahalanobis distance FCM-M produces a 5% increase. Finally, applying partial supervision adds another 4% so that the recognition rate becomes 9%.

Note that the low performance of the Zernike-IV is partly due to the individual selection. For some objects Zernike-AA-FCM-S could perform equally efficient.

The efficiency of the algorithm with the reference to the preceding techniques becomes significant when increasing the noise intensity (see Table 2-3). The most impressive result is an almost 29.4% absolute increase (37% relative increase) with regard to the Fourier Mellin-IV in the case of the musical instruments degraded by 3-4.5% impulse noise and the transformation noise (Table 2).

Table 1: Average classification rate of Thai musical instruments and NIST characters.

Algorithms	Instruments	NIST
	$0 \leq \eta < 4.5\%$	$0 \leq \eta < 6\%$
1.QMF-P-AA-FCM-S	93.24	95.61
2.QMF-P-AA-FCM-M	89.91	92.87
3.QMF-P-AA-FCM-E	84.74	89.91
3. Shen-AA-FCM-M	86.68	91.03
4. Shen-AA-FCM-E	83.94	89.17
5. Shen-IV	81.01	85.94
6. Zernike -IV	80.51	85.03
7. Fourier Mellin- IV	74.16	79.31

Table 2: Thai musical instruments, transformation noise and random-valued impulse noise:  $0 \leq \eta_1 < 1.5$ ,  $1.5 \leq \eta_2 < 3$  and  $3 \leq \eta_3 < 4.5$ .

Algorithm	$\eta_1$	$\eta_2$	$\eta_3$
QMF-P-AA-FCM-S	96.70	89.28	78.57
QMF-P-AA-FCM-M	92.84	80.18	67.18
Shen-AA-FCM-M	91.15	76.34	62.35
Shen- IV	89.82	74.52	58.51
Zernike -IV	89.88	70.19	56.96
Fourier Mellin-IV	86.17	65.32	49.13

Table 3: NIST database, transformation noise and random-valued impulse noise:  $0 \leq \eta_1 < 1.5$ ,  $1.5 \leq \eta_2 < 3$ ,  $3 \leq \eta_3 < 4.5$  and  $4.5 \leq \eta_4 < 6$ .

Algorithm	$\eta_1$	$\eta_2$	$\eta_3$	$\eta_4$
QMF-P-AA-FCM-S	98.91	94.47	88.40	80.52
QMF-P-AA-FCM-M	96.89	93.05	85.04	67.29
Shen-AA-FCM-M	96.30	90.92	81.19	60.93
Shen- IV	94.35	85.76	72.17	50.93
Zernike -IV	92.24	83.34	67.92	46.31
Fourier Mellin-IV	89.03	77.08	61.12	39.97

## 6. Conclusion

Our approach is designed for analysis of the rotationally invariant filter bank. The coefficients are analyzed using a combination of Fuzzy C-mean techniques and a modified data mining Apriori technique applied in supervised or unsupervised mode. The algorithm 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 5% for low noise, 15% for an

average noise and 20% 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.

## 7. Acknowledgement

We acknowledge a postdoctoral scholarship awarded to Dr A. Rodtook by the Thailand Research Fund (TRF) for this research.

## References

- [1] M. K. Hu, "Visual Pattern Recognition by Moment Invariants," *IRE Trans. Information Theory*, Vol. IT-8, pp. 179-187, 1962.
- [2] 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.
- [3] C. Kan and M. D. Srinath, "Invariant Character Recognition with Zernike and Orthogonal Fourier-Mellin Moments," *Pattern Recognition*, Vol. 35, pp. 143-154, 2002.
- [4] J. Flusser, "On the Inverse Problem of Rotation Moment Invariants," *Pattern Recognition*, Vol. 35, pp. 3015-3017, 2002.
- [5] D. Shen and H. H. Ip, "Discriminative Wavelet Shape Descriptors for Recognition of 2-D patterns," *Pattern Recognition*, Vol. 32, pp. 151-165, 1999.
- [6] S. Rodtook and S.S. Makhanov, "Numerical Experiments on Accuracy of Rotation Moment Invariants," *International Journal of Image and Vision Computing*, Vol. 23, pp. 577-586, 2005.
- [7] S.G. Mallat, "Multifrequency Channel Decomposition of Images and Wavelet Models," *IEEE Transactions on Image Processing*, Vol. 37, No. 12, pp. 2091-2110, 1989.
- [8] M. Unser and A. Aldroubi, "Polynomial Splines and Wavelets. A Signal Processing Perspective," in: *Wavelets- A Tutorial in Theory and Application*, Academic Press, Boston, pp. 91-122, 1992.
- [9] S. Abe, "Optimizing Features," in: *Pattern Classification*, Springer, London, UK, pp. 215-224, 2001.
- [10] M. Thuillard, "Preprocessing the Multiresolution Approach," in: *Wavelets in Soft Computing Series in Robotics and Intelligent Systems*, Vol. 25, World Science, pp.33-55, 2001.

- [11] R.A. Johnson, "Analysis of Variance," in: *Sixth Edition Probability and Statistic for Engineers*, Prentice Hall, New Jersey, pp. 407-414, 2000.
- [12] J. Han and M. Kamber, "Mining Association Rules in large Database," in: *Data Mining Concepts and Techniques*, Morgan Kaufmann Publishers, Academic Press, London, pp. 114-116, 2001.
- [13] The Fine Arts Departments of Thailand, "Visual Database of Thai Musical Instruments," *Manual of the Fine Arts Department*, Wattanapanit Press, Bangkok, Thailand, 2000.
- [14] National Institute of Standards and Technology, "Database 8 NIST Machine-Print DB of Gray Scale and Binary Images (MPDB), Standard Reference Data Program," in: <http://www.nist.gov/srd/nistsd8.htm>.
- [15] H.J. Sun, S.R. Wang and Z. Mei, "A Fuzzy Clustering Based Algorithm for Feature Selection," In: *Proceedings of the 1<sup>st</sup> International Conference Machine Learning and Cybernetics*, Beijin, China, pp. 1993-1997, 2002.
- [16] P.J. Deer, P.W Eklund and B.D. Norman, "A Mahalanobis Distance Fuzzy Classifier," in: *International conference of Intelligent Information Systems*, Australia, Vol. 1, pp. 220–223, 1996.
- [17] W. Pedrycz and J. Waletzky, "Fuzzy Clustering with Partial Supervision," *IEEE Trans. On Systems Man and Cybernetics*, Vol. 27, No. 5, pp. 787-795, 1997.
- [18] C.S. Burrus, R.A. Gopinath and H. Guo, *Introduction to wavelets and Wavelets Transforms*, Prentice Hall, New Jersey, US, 1998
- [19] S. Belongie, J. Malik and J. Puzicha, Shape Matching and Object Recognition Using Shape Contexts, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol 24, No 24, pp. 509-522, 2002