Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 28 (2007) 1492–1500
Pattern Recognition Letters
www.elsevier.com/locate/patrec

A filter bank method to construct rotationally invariant moments for pattern recognition

Annupan Rodtook [a,*], Stanislav S. Makhanov [b]

[a] Department of Computer Science, Ramkhamhaeng University, Bangkok 10240, Thailand
[b] School of Information and Computer Technology, International Institute of Technology, Thammasat University, Bangkadi Campus, Pathumthani 12000, Thailand

Received 27 June 2006; received in revised form 12 January 2007
Available online 20 March 2007

Communicated by G. Borgefors

Abstract

We propose multiresolution filter bank techniques to construct rotationally invariant moments. The multiresolution pyramid motivates a simple but efficient feature selection procedure based on a combination of a pruning algorithm, a new version of the Apriori mining techniques and the partially supervised fuzzy *C*-mean clustering. The recognition accuracy of the proposed techniques has been tested with the reference to conventional methods. The numerical experiments, with more than 50,000 images taken from standard image datasets, demonstrate an accuracy increase ranging from 5% to 27% depending on the noise level.

Keywords: Rotationally invariant moments; Wavelet filter bank; Feature selection; The Kullback–Leibler distance; Apriori mining algorithm; Fuzzy *C*-mean clustering

1. Introduction

The rotationally invariant moments are based on a spatial-frequency domain representation. A circular Fourier transform (the Fourier transform with regard to the angular coordinate) is applied inside a circle occluding the object. The result is a complex function the frequency and the radial variable. Next, the function is sampled with regard to the frequency variable at integer frequencies. The magnitude of every sample is rotationally invariant. Next, the samples, which are functions of the radial variable, are represented in terms of an appropriate basis. The coefficients of the Fourier series in this basis constitute the required moments. In this paper we will call these moments "coefficients" ("details" and "approximations") and "features" depending on the context. The choice of the basis is critical for pattern recognition. The most popular options are the Legendre polynomials, the Zernike polynomials, Mellin monomials (Teh and Chin, 1988; Mukundan and Ramakrishnan, 1995; Kan and Srinath, 2002) complex monomials (Flusser, 2002), the Tchebichef (Mukundan et al., 2001) and the Krawtchouk polynomials (Yap et al., 2003) (the Legendre, Tchebichef and Krawtchouk polynomials have been applied only in the Cartesian domain). Shen and Ip (1999) 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 "l", etc. It was confirmed by experiments with the accuracy of the wavelet rotation invariants published in (Rodtook and Makhanov, 2005). However, Shen and Ip do not fully exploit the concept of multiresolution. The wavelet moments are obtained by

* Corresponding author. Tel.: +66 2501 3505 30; fax: +66 2501 3524.
E-mail address: annupan@ru.ac.th (A. Rodtook).

0167-8655/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2007.03.003

integrating the circular Fourier transform of the object image with the projections onto the so-called Gaussian approximation of a B-spline the mother wavelet at different resolution levels. In the framework of the multiresolution analysis these integrals constitute the "details" coefficients suitable for recognition of similar objects yet belonging to different classes such as digit 1 and letter "l", etc. However, such recognition may fail when objects from the same class are subjected to random noise. It may also fail when the object is obtained from the original object by adding a single or several parts or making holes As opposed to that, our technique uses both the approximation and the detail coefficients of the multiresolution pyramid (filter bank). The multiresolution analysis treats the above cases efficiently and has a better recognition rate as compared with the preceding methods. Next, we propose a feature selection algorithm designed specifically for the multiresolution pyramid. First, it processes specific wavelet bands, then the wavelet coefficients individually, and, finally, combinations of the coefficients.

First, the circular Fourier transform of the object is sampled at integer frequencies called the angular orders. Next, each sample is subjected to a fast quadrature mirror filter (QMF) to generate a filter bank. The filter bank is characterized by a large number of coefficients and is always overcomplete. Many coefficients do not contribute and even degrade the performance of the classifier. Our new multi stage feature selection algorithm eliminates noise sensitive, redundant and non-important features. As opposed to Shen and Ip, the algorithm takes a full advantage of the multiresolution analysis. First of all, we exclude noise-sensitive frequencies. Next, we obtain a tree structured filter bank and prune it using the Kullback–Leibler distance which measures the relative entropy of the decompositions (Coifman and Wickerhauser, 1992). Next, we analyze the features individually by the standard ANOVA and feed the result to a selection procedure based on the Apriori Algorithm (AA) (Han and Kamber, 2001) initially developed for data mining in large databases. We propose a modified version of the AA (MAA) combined with a partially supervised fuzzy C-mean clustering technique (FCM) (Pedrycz and Waletzky, 1997). The FCM cost function is used in the MAA instead of a probabilistic measure employed by the standard AA to evaluate the confidence in combinations of the features. Furthermore, the conventional AA requires that combinations of features obey the so-called anti-monotonic property, that is, if a set cannot pass a test then all of its supersets fail the test as well. We propose a Δ anti-monotonic property which states that the combination can fail the test within a certain interval but still be considered at the next stage. Relaxing the conditions of the conventional AA allows to pass a local minimum and to find a better combination of the features. We show that the MAA combined with the partially supervised FCM performs extremely well on the multiresolution coefficients and creates features leading to high recognition rates. The recognition rate of the new algorithm has been tested by 50,000 different images and compared with the Zernike moments, the Fourier–Mellin moments as well as with wavelet based moments employing only details proposed by Shen and Ip. The algorithm has been also compared with a variety of previously reported feature selection techniques such as the individual selection, the selection which employs all combinations of the features and a concatenation of the MAA and the unsupervised FCM. We also analyze the use of the Euclidian and the Mahalonobis distance in the multiresolution framework.

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

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

Let $F(r,\theta) = \beta(r)\omega(\theta)$, 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. $\omega(\theta)$ denotes an angular function. Taking $\omega(\theta) \equiv \omega_q(\theta) = \mathrm{e}^{\mathrm{i}q\theta}$ provides the rotational invariance. Note, that if q is a continuous variable, then the integral with regard to θ is nothing but the circular Fourier transform. In the theory of rotationally invariant moments q is an integer called the angular order (Shen and Ip, 1999).We present the above 2-D integral by

$$M_q = \int_0^1 \beta(r)\xi_q(r)r\,\mathrm{d}r, \tag{2}$$

where

$$\xi_q(r) = \int_0^{2\pi} f(r,\theta)\mathrm{e}^{\mathrm{i}q\theta}\,\mathrm{d}\theta \tag{3}$$

is a complex function of the spatial variable r.

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

3. Rotationally invariant filter bank

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 an appropriate functional space. In the case of wavelets, the wavelet basis function is given by $\psi_{m,n}(r) = 2^{m/2}\psi(2^m r - 0.5n)$ (Daubechies, 1992), 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 moments will be introduced next.

In terms of the multiresolution analysis the sequence of approximating spaces is generated by the so-called scaling

functions φ (Daubechies, 1992; Strang, 1989; He, 2004), 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:

$$A_{m,n,q} = \int_0^1 \varphi_{m,n}(r)\xi_q(r) r\,\mathrm{d}r, \quad D_{m,n,q} = \int_0^1 \psi_{m,n}(r)\xi_q(r) r\,\mathrm{d}r, \tag{4}$$

where $\varphi_{m,n}(r)$ are the so-called scaling functions.

Note that $\sum_n A_{m+1,n,q}\varphi_{m+1,n}(r) = \sum_n D_{m,n,q}\psi_{m,n}(r) + \sum_n A_{m,n,q}\varphi_{m,n}(r)$ which tells us that ξ_q in a fine space $m+1$ is the sum of its representation in the coarse space m and its orthogonal complement. There exists a variety of the wavelet bases. A classic example is the Haar basis. The Haar wavelet and scaling function are given respectively by

$$\psi(r) = \begin{cases} 1, & t \in [0, 0.5) \\ -1, & t \in [0.5, 1) \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad \varphi(r) = \begin{cases} 1, & r \in [0,1) \\ 0, & \text{otherwise} \end{cases}.$$

However, in the general case a closed form for $\varphi(r)$ may not exist. In this case the scaling functions are to be found from the so-called dilation equation given by $\varphi_{m+1}(r) = \sum_{k=0}^{M_C} c_k \varphi_m(2r-k)$, where $\varphi_0(r) = \varphi(r)$ and $c_0, c_1, \ldots, c_{M_C}$ are the coefficients satisfying certain conditions (He, 2004). There are several ways to construct the scaling functions (Strang, 1989; Meyer, 1992). The Daubechies method (Daubechies, 1992) constructs the scaling functions by iterating the dilation equation above. Once the scaling functions are found, the wavelet function is represented by $\psi_{m+1}(r) = \sum_{k=0}^{M_C} (-1)^k c_k \varphi_m(2r-k)$. For instance for $M_C = 1$ and $M_C = 3$ the corresponding functions are, respectively, the Haar wavelet and the Daubechies D_4 wavelet. The discrete version of the above decomposition called the discrete wavelet multiresolution pyramid or the filter bank was proposed by Mallat (1989), Unser and Aldroubi (1992). It was shown that the discrete wavelet transform can be performed by using the so-called finite impulse response filters (FIRF) which produce a tree structured filter bank. This approach can be derived with or without a reference to the continuous version of the wavelets. In the framework of the QMF, the approximation and the detail filter bank moments are constructed using convolution type formulas as follows:

$$a_{m,n,q} = \sum_{\substack{k=0 \\ k-2n \geqslant 0}}^{N_{m+1}-1} h_{k-2n} a_{m+1,k,q}, \quad d_{m,n,q} = \sum_{\substack{k=0 \\ k-2n \geqslant 0}}^{N_{m+1}-1} g_{k-2n} a_{m+1,k,q}. \tag{5}$$

$m = m_0, m_0-1, \ldots, 0$ is the scale (resolution) index, $n = 0, 1, \ldots, (N_{m+1}/2) - 1$ is the position index. Each step reduces the number of available coefficients N_{m+1} by factor 2. $m_0 + 1$ is the highest resolution level. The input $a_{m_0+1,n,q}$ is obtained by sampling (3) at $r_n = n/(N-1)$, $n = 0, \ldots, N-1$, where $N = N_{m_0+1}$ is the size of the input data.

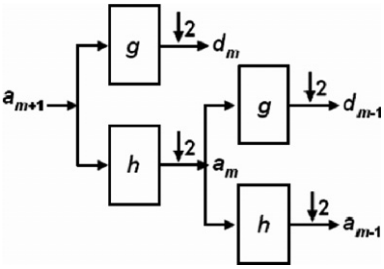

Fig. 1. A two-band filter bank tree. Variables q and n have been omitted. Only 2 steps are shown.

Finally, h and g are the low-pass and the high pass FIRFs associated with a corresponding wavelet. The popular Daubechies and B-spline based FIRFs of an arbitrary order can be obtained from Matlab 7 or from the Wavelet Explorer of Mathematica. Note that in (5) the data is extended periodically at the end points using a procedure similar to the circular convolution. Then the output is downsampled and a low-frequency output (approximation) is fed to the identical filters as shown in Fig. 1.

The number of coefficients out of this system is the same as the number in. The number is doubled by having two filters; then it is halved by the decimation back to the original number. Actually, no information is lost in this scheme and it is possible to completely recover the original coefficients. Furthermore, let $\tilde{f}$ be a rotated image, where ϕ is the angle of rotation, then

$$\tilde{a}_{m,n,q} = \sum_k h_{k-2n}\tilde{a}_{m+1,k,q} = \mathrm{e}^{\mathrm{i}q\phi} \sum_k h_{k-2n} a_{m+1,k,q},$$
$$\tilde{d}_{m,n,q} = \sum_k g_{k-2n}\tilde{a}_{m+1,k,q} = \mathrm{e}^{\mathrm{i}q\phi} \sum_k g_{k-2n} a_{m+1,k,q}.$$

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

4. Feature selection algorithm

We present an algorithm based on examining the bands and the features individually and in combinations. As opposed to (Shen and Ip, 1999) our algorithm takes full advantage of the multiresolution analysis. First, we prune the filter bank using the relative entropy of the decompositions. Next, we analyze the bands and contributions of each individual wavelet as well. We test the features individually by the standard ANOVA and feed the result to a selection procedure based on a new modification of the Apriori technique (MAA) to analyze combinations of the coefficients. The algorithm is given below:

1. ***Calculate*** $\xi_q(r)$ ***for a prescribed set of the angular orders*** q by sampling the circular Fourier transform (3).
2. ***Discard noise-sensitive angular orders*** by evaluating the least square error type given by

$$E(q) = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{n=0}^{N-1}\left(\left|\xi_q(r_n)^{i,\text{Template}}\right| - \left|\xi_q(r_n)^{i,j}\right|\right)^2}{IJN}, \tag{6}$$

where I is the number of the classes, J the number of objects in each class and $\xi_q(r_n)^{i,\text{Template}}$ the circular Fourier transform of the template associated with class i. Note that the procedure is just a preliminary filtering which allows to discard easily recognizable bad choices. The resulting set is fed to the next step of the procedure.

3. ***Apply the QMF to the output of step 2*** to obtain a tree structured filter bank of coefficients (approximation and detail moments).
4. ***Prune the resulting filter bank*** to find the best discriminating subbands which produce well-separated classes (Coifman and Wickerhauser, 1992; Saito and Coifman, 1994). A symmetric version of the Kullback–Leibler distance based on the relative entropy is used to measure the discrimination power of the subband. A good subband is the one that reduces the relative entropy, which for the case of two classes C_1 and C_2 is given by

$$\delta_{C_1C_2}(m,q,k) = \frac{1}{2}\sum_{n=0}^{N_m-1}\left(\gamma_{C_1}(m,q,n,k)\log\frac{\gamma_{C_1}(m,q,n,k)}{\gamma_{C_2}(m,q,n,k)} + \gamma_{C_2}(m,q,n,k)\log\frac{\gamma_{C_2}(m,q,n,k)}{\gamma_{C_1}(m,q,n,k)}\right), \tag{7}$$

where

$$\gamma_{C_i}(m,q,n,k) = \frac{\sum_{w_{m,n,q}\in C_i} w_{m,n,q}w^*_{m,n,q}}{\sum_{a_{m+1,n,q}\in C_i} a_{m+1,n,q}a^*_{m+1,n,q}}.$$

$*$ denotes complex conjugate and $k=0$, 1 denotes the approximation and the detail branch of the tree respectively, and the summation is over all objects from class C_i

$$w_{m,n,q} = \begin{cases} a_{m,n,q}, & \text{if } k=0 \text{ (approximations)} \\ d_{m,n,q}, & \text{if } k=1 \text{ (details)}. \end{cases} \tag{8}$$

In the case of I classes

$$\delta(m,q,k) = \sum_{i=1}^{I-1}\sum_{j=i+1}^{I}\delta_{C_iC_j}(m,q,k). \tag{9}$$

Using (7)–(9), the algorithm evaluates the discrimination power of the subbands by comparing the Kullback–Leibler distance before and after the split as shown in Fig. 2.

5. ***Reduce the dimension of the feature space*** by analyzing the features individually using a statistical testing (Johnson, 2000). We use a one-way ANOVA applied to the magnitudes of the moments with a randomized complete block design to verify the assumption $\mu_1 \neq \mu_2 \neq \cdots \neq \mu_i \cdots \neq \mu_I$, where μ_i is the mean-feature of the class i.
6. ***Analyze combinations of the features.*** At this stage the result of the multiresolution analysis is fed to the Modified Apriori Algorithm (MAA). The Apriori selection initially developed for data mining applications reduces the number of combinations appearing when mining for frequent itemsets in large databases. Since the filter bank produces a large number of features, the AA is beneficial in this case as well. An initial set of the "frequent" features L_1 is found by ANOVA at step 5. It is then used to find L_2 which consists of the best pairs of features taken from L_1. The set of the best discriminating 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 $C_k = L_{k-1}\times L_{k-1}$. The choice of a good combination of the features is based on evaluating a partially supervised FCM-type cost function (Pedrycz and Waletzky, 1997) used by analogy with the measure of confidence in the conventional Apriori algorithm (Han and Kamber, 2001). The cost function is given by

$$f_C(X) = J(X)\log(N_{\text{miss}} + \vartheta), \tag{10}$$

where X is a combination of features, J is the fuzzy C-mean cost function, N_{miss} is the number of training patterns that have been incorrectly clustered, $\vartheta > 1$ a prescribed constant to eliminate the singularity log (0). Besides the features should be normalized or standardized prior to calculating f_C. Furthermore, the conventional AA requires that combinations of features obey the so-called anti-monotonic property, that is, if a set cannot pass a test then all of its supersets fail the test as well. For example, in the case of two features A and B the anti-monotonic property requires that

$$f_C(A\cup B) \leqslant f_C(A) \quad \text{and} \quad f_C(A\cup B) \leqslant f_C(B). \tag{11}$$

As opposed to that, the MAA employs a Δ-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

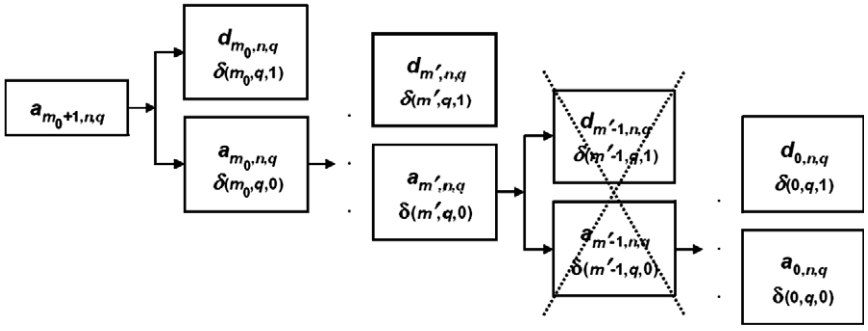

Fig. 2. An example of pruning by the Kullback–Leibler relative entropy. Band $m'-1$ has been removed since $\delta(m',q,0) < \delta(m'-1,q,0) + \delta(m'-1,q,1)$.

B with regard to a cost function f_C ($f_C \geqslant 0$) is evaluated as follows:

$$f_C(A \cup B) \leqslant f_C(A) + \Delta \quad \text{and} \quad f_C(A \cup B) \leqslant f_C(B) + \Delta, \tag{12}$$

where Δ defines the allowable interval of confidence. The confidence in a combination of n features is evaluated using the same principle. The best combination is the k-itemset which minimizes the cost function above.

Finally, the MAA applies as follows: (1) find the best feature set from each particular resolution, (2) find the best feature set from the entire multiresolution analysis, (3) find the best feature set from the entire set of the angular orders.

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

5. Why does the feature selection perform better? Illustrative examples

This section exemplifies some steps of the algorithm described above applied to differentiate between digit 1 and letter "1".

Example 1 (*Pruning the Filter Bank Tree*). Fig. 3 illustrates the pruning step for the filter bank tree created for the angular order at $q = 1$ for recognition between digit 1 and letter "1". The coarsest level $m = 0$ as well as level $m = 2$ have been eliminated since $\delta(1,0,1) = 0.082 < \delta(0,0,1) + \delta(0,1,1) = 0.098$ and $\delta(3,0,1) = 0.394 < \delta(2,0,1) + \delta(2,1,1) = 0.423$. Note, that although level $m = 2$ has been excluded, it does not mean that all the subsequent levels are not appropriate. For example, level $m = 1$ has also been selected.

Finally, the pruning algorithm combined with the individual statistical selection by ANOVA has eliminated 47.4% of the features (74 from 156 available coefficients), thus, substantially reducing the number of combination to be tested at step 5.

Example 2 (*Modified Apriori Algorithm*). The MAA applied to feature selection for differentiation between 1 and "1" is illustrated here. We check all combinations taken from the original set at the resolution level $m = 3$ consisting of 40 coefficients (20 details and 20 approximations). Consider how the MAA applies to this case. The first itemset L_1 is an output of ANOVA. It consists of 12 approximation and 13 detail coefficients. The set is given by

$$\{a_2\},\{a_3\},\{a_4\},\{a_5\},\{a_6\},\{a_{10}\},\{a_{12}\},\{a_{15}\},\{a_{16}\},\{a_{17}\},\{a_{18}\},\{a_{19}\},\{d_1\},\{d_3\},\{d_6\},\{d_8\},\{d_9\},\{d_{10}\},\{d_{12}\},\{d_{13}\},\{d_{14}\},\{d_{15}\},\{d_{16}\},\{d_{17}\},\{d_{18}\}.$$

The next level L_2 contains only 16 items:

$$\{a_3,a_5\},\{a_3,a_6\},\{a_3,d_9\},\{a_5,a_6\},\{a_5,d_9\},\{a_6,d_9\},\{a_{15},a_{16}\},\{a_{15},d_8\},\{a_{15},d_{14}\},\{a_{15},d_{17}\},\{a_{16},d_8\},\{a_{16},d_{14}\},\{a_{16},d_{17}\},\{d_8,d_{14}\},\{d_8,d_{17}\},\{d_{14},d_{17}\}.$$

L_3 has 11 items:

$$\{a_3,a_5,d_9\},\ \{a_{15},a_{16},d_8\},\ \{a_{15},a_{16},d_{14}\},\ \{a_{15},d_{16},d_{17}\},\ \{a_{15},d_8,d_{14}\},\ \{a_{15},d_8,d_{17}\},\ \{a_{15},d_{14},d_{17}\},\ \{a_{16},d_8,d_{14}\},\ \{a_{16},d_8,d_{17}\},\ \{a_{16},d_{14},d_{17}\},\ \{d_8,d_{14},d_{17}\}.$$

Finally L_4 includes only 1 item: $\{a_{15},a_{16},d_{14},d_{17}\}$.

Note that although a_k, d_k complex numbers, evaluation of the cost function f_C requires only their (rotationally invariant) magnitudes.

Suppose that the set of the best combinations of features is selected using the condition $f_C \leqslant 17$. In this case our data mining algorithm outputs: $\{a_{15},d_{14}\}$, $\{a_{15},d_{17}\}$, $\{d_{14},d_{17}\}$, $\{a_{15},d_{14},d_{17}\}$ and $\{a_{15},a_{16},d_{14},d_{17}\}$. The best combination is $\{a_{15},d_{14},d_{17}\}$, $f_C(\{a_{15},d_{14},d_{17}\}) = 15.9086$. However, the conventional AA would discard $\{a_{15},a_{16},d_{14}\}$ since $f_C(\{a_{15},a_{16},d_{14}\}) > f_C(\{a_{15},d_{14}\})$. Consequently a good combination $\{a_{15},a_{16},d_{14},d_{17}\}$ which satisfies f_C ($\{a_{15},a_{16},d_{14},d_{17}\}$) $= 16.8072 \leqslant 17$ is not selected. However, the MAA characterized by $\Delta = 2$ will select $\{a_{15},a_{16},d_{14},d_{17}\}$. Note, that the case of $20 + 20$ coefficients above was considered for simplicity. Taking, for example, $32 + 32$ coefficients leads to astronomical 2×10^{10} Gflops which for a desktop computer with a Pentium 4 means approximately 300 years of calculations. However, the proposed method requires only a few hours. The choice of Δ depends on the number of combinations this choice generates. Of course, ideally, we would like to check all possible combinations. However, more often than not this leads to an

Fig. 3. Pruning subbands at resolution level $m = 0$ and $m = 2$ by using the Kullback–Leibler distance.

astronomical time of computation. Therefore, Δ should be set in such a way that the computing time is reasonable. For example, one can consider 10–15% of the minimum of cost function at each step. Actually, it is also possible to threshold the maximum allowed number of combinations. However, in this case combinations with a poor discrimination power could be selected. Finally, since the number of coefficients is decreasing at each step, it is also possible to vary Δ depending on the step.

6. Experimental results

We evaluate the performance of the proposed algorithm by three datasets. The first dataset consists of 48,400 noisy images based on twenty basic aircraft silhouettes (Wickham, 1986): Alpha Jet, Am-X, Mirage F1, F-4 Phantom, F-15 Eagle, MiG-17, A-6 Intruder, Aviocar C-212, An-32 Cline, F-5 Freedom, An-12 Cub, Hunter, Brewer, Jastreb, MiG-29, Buccaneer, MiG-25, Mirage III, F-18 Hornet, Yak-36 (see Fig. 4). Each silhouette produces 1600 training images and 820 testing images. The second dataset based on an online database (NIST database), consists of machine-printed characters, namely, 11000 upper case English letters (Bold, Courier). We use 7000 letters for training and 4000 for testing. The third dataset consists of 8640 images based on sixteen corporate 8-bit logo images (UMD Database) taken from a research database of University of Maryland (UMD), see Fig. 5. Each logo produces 351 training images and 189 testing images. Since the UMD logos are very different, the images were modified in order to create a more difficult pattern recognition task. For each image we created a "similar" image by flipping a part of the picture as shown in Fig. 5. For example, "Kodak" in image (c) was replaced by "kadoK" in (d), 66 in image (i) was changed to 99 in image (j), etc., so that the images have identical histograms but still are different. It is of interest to note that flipping the entire image does not change the circular Fourier transform. Therefore, the method is not able to differentiate between the flipped and the original image. Furthermore, all of the dataset images have been degraded by an impulse noise varying from 0% to 8% and a transformation (rotation and scaling)

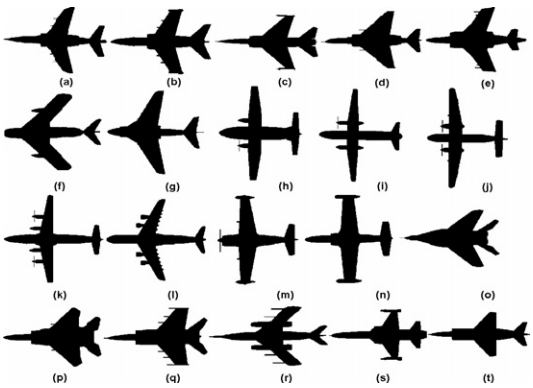

Fig. 4. Twenty silhouettes of Aircraft.

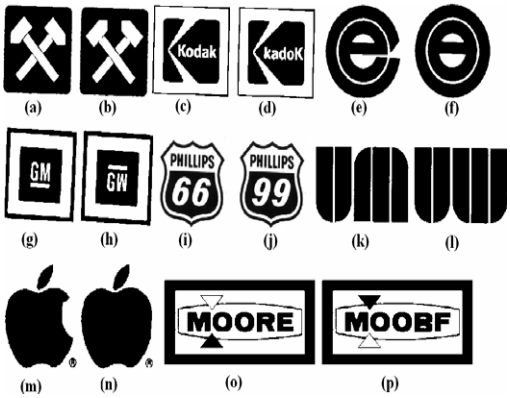

Fig. 5. The modified UMD logos.

noise (Rodtook and Makhanov, 2005). Additionally, in the case of the NIST data we also considered an interesting effect of the boundary noise appearing after separation of touching letters by means of morphological noise removal employing two erosions and one dilation. We present experiments with the B-spline wavelet filters, however, orthogonal wavelets such as the Daubechies wavelets 2, 4 and 6 and the Coiflet wavelets were tested as well. The experiments have shown that the biorthogonal B-splines with the underlying FIRFs were always performing slightly better. Denote our proposed algorithm by FB-P-AN-MAA-FCM-P, FB stands for the proposed filter bank, P for pruning, AN for ANOVA, MAA for the modified Apriori algorithm, FCM for Fuzzy C mean clustering, P for partially supervised FCM, so that FCM-P corresponds to the FCM with cost function (10). Furthermore, FCM-E and FCM-M correspond to the unsupervised FCM endowed with the Euclidian and the Mahalanobis distance respectively. Finally, the notation I–V is used when the features were selected individually based on the between- to within-class variance ratio. The comparisons of an average classification rate of the proposed FB-P-AN-MAA-FCM-P versus the most popular selection methods and moment invariants are shown in Table 1. Table 1 includes degradation by all types of noise: rotation, translation, scaling and a random impulse noise η.

Table 1 illustrates advantages of our approach. For instance, "Shen-I–V" applied to the NIST symbols has 85.94% average recognition rate, whereas our method provides recognition rate of about 95.46%. The table shows that every component of the algorithm is almost equally important. Namely, combining the FB with the FCM shows a 3% increase. Adding the Mahalanobis distance produces a 6% increase. Finally, applying partial supervision adds another 3% so that the recognition rate becomes improved by 9%.

Tables 2–7 show the average classification rate for different type of image distortions such as the random impulse noise, transformation noise, segmentation noise, scaling as well as certain combinations of them. In each case the algorithm outperforms Shen-IV, Shen-FCM-M, as well as Zernike-IV and Fourier–Mellin-IV. The efficiency of the

Table 1
Average classification rate within the three databases

Algorithms	Classification rates %		
	Aircraft silhouettes $0 \leqslant \eta \leqslant 8\%$	NIST characters $0 \leqslant \eta \leqslant 6\%$	Logo silhouettes $0 \leqslant \eta \leqslant 4.5\%$
1. FB-P-AN-MAA-FCM-P	95.38	95.46	92.81
2. FB-P-AN-MAA-FCM-M	91.19	92.76	88.72
3. FB-P-AN-MAA-FCM-E	88.05	89.73	84.30
3. Shen-FCM-M	88.51	90.94	85.76
4. Shen-FCM-E	85.17	89.1	83.01
5. Shen-I–V	82.94	85.94	79.12
6. Zernike-I–V	82.19	85.03	78.93
7. Fourier–Mellin-I–V	77.61	79.31	73.72

Table 2
Aircraft images, impulse noise

Algorithm	Noise % SNR (dB)	$0 \leqslant \eta < 2$ >15.9	$2 \leqslant \eta < 4$ >12.5	$4 \leqslant \eta < 6$ >10.4	$6 \leqslant \eta < 8$ >9.0
FB-P-AN-MAA-FCM-P		98.83	95.49	91.52	81.21
FB-P-AN-MAA-FCM-M		97.71	92.91	85.13	67.09
Shen-FCM-M		96.28	90.39	81.16	61.31
Shen-I–V		95.89	85.35	72.39	53.49
Zernike-I–V		95.52	84.33	70.91	51.41
Fourier–Mellin-I–V		90.24	77.78	59.43	43.62

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

Algorithm	Noise % SNR (dB)	$0 \leqslant \eta < 2$ >15.9	$2 \leqslant \eta < 4$ >12.5	$4 \leqslant \eta < 6$ >10.4	$6 \leqslant \eta < 8$ >9.0
FB-P-AN-MAA-FCM-P		98.45	94.66	90.28	78.53
FB-P-AN-MAA-FCM-M		95.76	91.08	81.83	62.26
Shen-FCM-M		95.01	89.84	78.02	54.25
Shen-I–V		92.88	82.3	65.36	45.01
Zernike-I–V		92.24	80.36	62.43	41.26
Fourier–Mellin-I–V		87.76	71.43	47.01	34.74

Table 4
The NIST characters, impulse noise and segmentation noise

Algorithm	Noise % SNR (dB)	$0 \leqslant \eta < 1.5$ >16.7	$1.5 \leqslant \eta < 3$ >14.0	$3 \leqslant \eta < 4.5$ >12.1	$4.5 \leqslant \eta < 6$ >10.7
FB-P-AN-MAA-FCM-P		99.01	95.83	90.85	86.23
FB-P-AN-MAA-FCM-M		98.97	93.92	87.98	72.12
Shen-FCM-M		98.12	91.82	84.83	67.98
Shen-I–V		97.04	88.65	80.57	61.05
Zernike-I–V		96.9	87.73	78.03	58.24
Fourier–Mellin-I–V		94.27	81.82	70.12	49.61

algorithm with the reference to the preceding techniques becomes significant when increasing the noise intensity. The most impressive result is an almost 44% absolute increase (55% relative increase) with regard the Fourier–Mellin-IV in the case of the aircraft silhouettes degraded by 6–8% impulse noise and the rotation noise (Table 2).

The rotation and segmentation noise affect the NIST characters (Table 5) more significantly than the aircraft silhouettes and the UMD logos (Tables 3 and 7) because the characters are thin objects with small area, so the relative impact of the noise is higher. Note, that we applied the Shen's algorithm exactly as it was presented in (Shen and Ip, 1999). The tables show that the algorithm will benefit from some steps we introduced as well. For example, in each table Shen-FCM-M outperforms Shen-I–V. Zernike-FCM-M will also display a similar accuracy increase. However, pruning is not applicable because Shen's invariants are not structured as a tree. The MAA can be applied to

Table 5
The NIST characters, impulse noise and transformation noise

Algorithm	Noise % SNR (dB)	$0 \leqslant \eta < 1.5$ >16.7	$1.5 \leqslant \eta < 3$ >14.0	$3 \leqslant \eta < 4.5$ >12.1	$4.5 \leqslant \eta < 6$ >10.7
FB-P-AN-MAA-FCM-P		98.85	94.4	88.34	80.51
FB-P-AN-MAA-FCM-M		96.83	92.96	84.98	67.24
Shen-FCM-M		96.21	90.84	81.15	60.87
Shen-I–V		94.35	85.76	72.17	50.93
Zernike-I–V		92.24	83.34	67.92	46.31
Fourier–Mellin-I–V		89.0	77.08	61.12	39.9

Table 6
Modified UMD logos, impulse noise

Algorithm	Noise % SNR (dB)	$0 \leqslant \eta < 1.5$ >17.7	$1.5 \leqslant \eta < 3$ >15.1	$3 \leqslant \eta < 4.5$ >13.2
FB-P-AN-MAA-FCM-P		96.13	90.78	82.57
FB-P-AN-MAA-FCM-M		94.06	85.71	74.70
Shen-FCM-M		92.87	84.91	72.85
Shen-I–V		91.09	80.34	69.07
Zernike-I–V		90.91	79.77	68.98
Fourier–Mellin-I–V		88.23	75.96	65.97

Table 7
Modified UMD logos, impulse noise and transformation noise

Algorithm	Noise % SNR (dB)	$0 \leqslant \eta < 1.5$ >17.7	$1.5 \leqslant \eta < 3$ >15.1	$3 \leqslant \eta < 4.5$ >13.2
FB-P-AN-MAA-FCM-P		95.49	89.02	77.10
FB-P-AN-MAA-FCM-M		92.61	79.14	66.91
Shen-FCM-M		89.87	74.81	62.14
Shen-I–V		87.06	72.18	58.03
Zernike-I–V		86.76	70.36	57.61
Fourier–Mellin-I–V		85.02	63.71	50.03

Shen' invariants, but it would not increase the recognition rate. It can only reduce the computation time. Observe the most interesting cases. In case of the NIST characters (see Table 5, 4.5–6% noise) and a combination of the impulse noise and the transformation noise Shen-IV, Zernike-IV and Fourier–Mellin-IV display 50.93%, 46.31% and 39.90% recognition rate respectively whereas the proposed filter bank invariants 80.51%. In case of the UMD logos and 3–4.5% random-valued impulse noise combined with the transformation noise the Shen-IV, Zernike-IV and Fourier–Mellin-IV display 58.03%, 57.61% and 50.03% recognition rate respectively whereas the proposed filter bank invariants show 77.10% recognition rate.

7. Conclusions

The proposed multiresolution moment 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 noise-sensitive features and the modified Apriori-Fuzzy *C*-mean partly supervised selection leads to a tangible improvement of the recognition rate with the reference to the conventional methods. We obtain an accuracy increase ranging from 5% to 27% depending on the noise level. 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.

Acknowledgement

We acknowledge a post graduate scholarship of the Thailand Research Fund granted to Dr. Annupan Rodtook to conduct this research.

We are also grateful to the referees of this paper for their valuable comments, suggestions and attention to the details.

References

Coifman, R.R., Wickerhauser, M.V., 1992. Entropy-based Algorithms for best basis selection. IEEE Trans. Inform. Theory 38 (2), 713–718.

Daubechies, I., 1992. Ten Lectures on Wavelets, CBMS-NSF Series in Applied Mathematics. SIAM, Philadelphia.

Flusser, J., 2002. On the independence of rotation moment invariants. Pattern Recognition 33 (9), 1405–1410.

Han, J., Kamber, M., 2001. Mining Association Rules in Large Databases. Data Mining. Academic Press, San Francisco, pp. 225–269.

He, Tian-Xiao, 2004. Biorthogonal spline type wavelets. Comput. Math. Appl. 48 (9), 1319–1334.

Johnson, R.A., 2000. Analysis of Variance. Sixth Edition Probability and Statistics for Engineers. Prentice-Hall, NJ, pp. 407–414.

Kan, C., Srinath, M.D., 2002. Invariant character recognition with Zernike and orthogonal Fourier–Mellin moments. Pattern Recognition 35 (1), 143–154.

Mallat, S.G., 1989. Multifrequency channel decompositions of images and wavelet models. IEEE Trans. Image Process. 37 (12), 2091–2110.

Meyer, Y., 1992. Wavelets and Operators. Cambridge University Press, Cambridge.

Mukundan, R., Ramakrishnan, K.R., 1995. Fast computation of Legendre and Zernike moments. Pattern Recognition 28 (9), 1433–1442.

Mukundan, R., Ong, S.H., Lee, P.A., 2001. Image analysis by Tchebichef moments. IEEE Trans. Image Process. 10 (9), 1357–1364.

NIST Database. National Institute of Standards and Technology. In: Database 8 NIST Machine-Print DB of Gray Scale and Binary Images (MPDB). <http://www.nist.gov/srd/nistsd8.htm>.

Pedrycz, W., Waletzky, J., 1997. Fuzzy clustering with partial supervision. IEEE Trans. Systems Man Cybernet. 27 (5), 787–795.

Rodtook, S., Makhanov, S.S., 2005. Numerical experiments on the accuracy of rotation moment invariants. Image Vision Comput. 23 (6), 577–586.

Saito, N., Coifman, R.R., 1994. Local Discriminant Bases. In: Proc. SPIE 2303 on Wavelet Applications in Signal and Image Processing II, San Diego, pp. 2–14.

Shen, D., Ip, H.H., 1999. Discriminative wavelet shape descriptors for recognition of 2-D patterns. Pattern Recognition 32 (2), 151–165.

Strang, G., 1989. Wavelets and dilation equations: A brief introduction. SIAM Rev. 31, 614–627.

Teh, C.H., Chin, R.T., 1988. On image analysis by the methods of moments. IEEE Trans. Pattern Anal. Machine Intell. 10 (4), 496–512.

UMD Database. University of Maryland Logo Database. <http://documents.cfar.umd.edu/resources/database/UMDlogo.html>.

Unser, M., Aldroubi, A., 1992. Polynomial splines and wavelets – a signal processing perspective. In: Wavelets. A Tutorial in Theory and Application. Academic Press, Boston, pp. 91–122.

Wickham Jr., J.A., 1986. Visual aircraft recognition. Manual of Headquarters Department of the Army. US Army Air Defense Artillery School, Washington DC.

Yap, P., Paramesran, R., Ong, S., 2003. Image analysis by Krawtchouk moments. IEEE Trans. Image Process. 12 (11), 1367–1377.