Available online at www.sciencedirect.com
ScienceDirect
Mathematics and Computers in Simulation 79 (2009) 2458–2475
MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

Selection of multiresolution rotationally invariant moments for image recognition

A. Rodtook [a,*], S.S. Makhanov [b]

[a] *Computer Science Department, Ramkhamhaeng University, Bangkok, Thailand*
[b] *Information Technology Program, School of Information, Computer and Communication Technology, Sirindhorn International Institute of Technology, Thammasat University, Pathumthani, Thailand*

Available online 3 February 2009

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 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, demonstrate an accuracy increase ranging from 5% to 27% depending on the noise level.

Keywords: Pattern recognition; Rotationally invariant moments; Wavelets; Filter bank

1. Introduction

A popular class of the rotationally invariant features is based on the moment techniques which are believed to be reliable for complex shapes because they involve not solely the contour pixels but all the pixels constituting the object. In 1962 Hu [12] published the first paper on the use of the image moments based on nonlinear combinations of central geometric moments invariant under translation and rotation (see also [17]). The Hu's technique has been used in many applications. However, a dramatic increase in complexity when increasing the order often makes Hu's moments impractical. Shortly after Hu's paper, a variety of rotationally invariant moments has been proposed and analyzed [5,6,15,16,21,20,22,27,29]. The main idea 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 inside a circle occluding the object. The result is a complex function of two real variables, 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 functional 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 moments are complex numbers, the magnitude of which is invariant with regard to the rotations of

* Corresponding author.
E-mail address: sittisak168@yahoo.com (A. Rodtook).

0378-4754/$36.00 © 2009 Published by Elsevier B.V. on behalf of IMACS.
doi:10.1016/j.matcom.2009.01.016

the object. Note that often it is not necessary to exactly separate the object from the background. It is sufficient to only find the occluding circle since the above procedure treats each pixel irrespectively whether it belongs to the object or not. Therefore, even in the case when the object boundary is not known, the rotation invariant moment technique is often applicable. However, pattern recognition problems with a very distorted background patterns may exist where this claim is not true.

Furthermore, the choice of the spatial basis could be critical for pattern recognition. The most popular options are the Legendre polynomials [21,29], the Zernike polynomials [15,16,21,29], Mellin monomials (the Fourier–Mellin moments) [15,29], complex monomials (complex moments) [5,6,29], the Tchebichef [20] and the Krawtchouk polynomials [22](the Legendre, the Tchebichef and the Krawtchouk polynomials have been applied only in the Cartesian domain).

Finally, Shen and Ip [27] introduced a set of wavelet moment invariants and a discriminative feature selection method for the classification of seemingly similar objects with subtle differences. The features were selected automatically based on the discrimination measures defined for the invariant features. Using a minimum-distance classifier, the wavelet moment invariants achieved the highest classification rate for four test sets with the reference to the Zernike's invariants and Hu's invariants.

Experiments with accuracy of the wavelet rotation invariants were published in [25].

However, [27] does not fully exploit the concept of multiresolution. The wavelet moments are obtained by integrating the circular Fourier transform of the object image with the projections onto the mother wavelet at different resolution levels. In the framework of the multiresolution analysis these integrals constitute the so called "details" suitable for recognition of similar objects yet belonging to different classes such as digit 1 and letter "l". 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 (see our forthcoming introductory and illustrative examples). As opposed to that, our technique uses both the approximation and the detail coefficients of the multiresolution pyramid (filter bank). We show that the multiresolution analysis combined with our new feature selection algorithm, designed specifically for the multiresolution pyramid, treats the above cases efficiently and has a better recognition rate as compared with the preceding methods.

First of all, the method processes specific wavelet bands, then the wavelet coefficients individually, and, finally, combinations of the coefficients. The circular Fourier transform of the object is sampled at integer frequencies called the angular orders. 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 multistage feature selection algorithm eliminates noise sensitive, redundant and non-important features. As opposed to [27] the algorithm takes a full advantage of the multiresolution analysis. First of all, we exclude noise sensitive frequencies. Next, we use a tree-structured filter bank and prune the tree using the Kullback–Leibler distance which measures the relative entropy of the decompositions. Next, we analyze the features individually by the standard ANOVA and feed the result to a selection procedure based on the Apriori technique. The Apriori algorithm (AA) initially developed for data mining reduces the number of combinations appearing when mining for frequent itemsets in large databases. Since the filter bank may produce a large number of features, application of the AA is beneficial in our case as well. We propose a modified version of the AA (MAA) combined with a partially supervised fuzzy C-mean clustering technique (FCM) [23,24]. 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 the so-called 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 conventional AA condition 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 the procedures and methods described in [27]. 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, concatenation of the AA and the unsupervised FCM. Finally, we analyze the use of the Euclidian and the Mahalonobis distance in the proposed framework.

The paper is organized as follows. Section 2 presents basics of rotationally invariant moments. We introduce the multiresolution filter-bank moments and analyze their efficiency in Sections 3 and 4, respectively. Section 5 deals with our approach to feature selection and classification and, in particular, with the development of our new feature selection algorithm. Section 6 demonstrates the performance of the proposed method applied to some binary images whereas Section 7 presents a variety of object recognition experiments on the gray level images and demonstrates the efficiency of the proposed schemes. Section 8 contains brief concluding remarks.

2. Basic definitions

2.1. The geometric moments

A regular moment $m_{p,q}$ is defined in the Cartesian coordinates by

$$m_{p,q} = \int\int x^p y^q f(x, y)\, dxdy, \tag{1}$$

where p, q are integer. In the context of image processing $f(x, y)$ is the image function representing the light intensity (the gray level) of the object. Furthermore, the moments can be also defined in the polar coordinate system with the origin at the centroid of the object image given by

$$x_C = \frac{m_{1,0}}{m_{0,0}}, \qquad y_C \frac{m_{0,1}}{m_{0,0}}.$$

2.2. 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

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

Let $F(r, \theta) = \beta(r)\omega(\theta)$, where $\beta(r) \in \Re$ and $\Re$ is a family of radial functions such as the Zernike polynomials [15,16,21,29], Mellin polynomials [15,29], etc. The choice of $\Re$ defines the type of the moment. Furthermore, $\omega(\theta)$ denotes an angular function. Taking $\omega(\theta) \equiv \omega_q(\theta) = e^{iq\theta}$ provides the rotational invariance. Note, that if q is a continuous variable, then the integral with regard to θ is nothing but the circular Fourier transform. In the theory of rotationally invariant moments q is an integer called the angular order [27]. We present the above 2D integral by

$$M_q = \int_0^1 \beta(r)\xi_q(r)rdr, \tag{3}$$

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 ϕ 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.

Finally, note that $\int_0^1 \beta(r)|\xi_q(r)|rdr$ and $\int_0^1 |\beta(r)\xi_q(r)rdr|$ are rotationally invariant as well.

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 $\Re$ constitutes a basis in the appropriate functional space. In the case of wavelets, $\Re$ consists

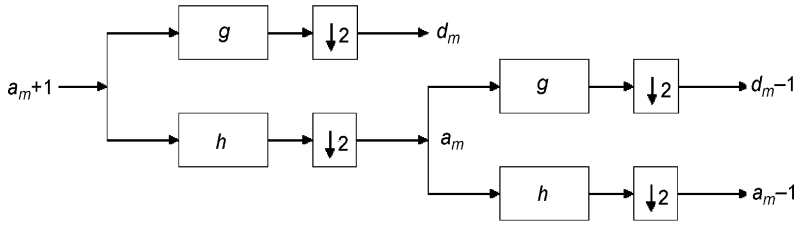

Fig. 1. Two-resolution two-band filter bank tree.

of the wavelet basis functions given by $\psi_{m,n}(r) = 2^{m/2}\psi(2^m r - 0.5n)$ [27], 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 [3,18] 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:

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

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 called the discrete wavelet multiresolution pyramid or the filter bank was proposed by Mallat [3,18]. 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. This approach can be derived with or without a reference to the continuous version of the filter bank. Mallat proposed a quadrature mirror filter which corresponds to the orthogonal wavelets. Unser and Aldroubi [33] extended these techniques to the biorthogonal wavelets.

Within the framework of the QMF concept, the approximation and detail filter bank moments are constructed as follows:

$$a_{m,n,q} = \sum\nolimits_k h_{k-2n} a_{m+1,k,q}, \qquad d_{m,n,q} = \sum\nolimits_k g_{k-2n} a_{m+1,k,q}, \tag{4}$$

$m = m_0, m_{0-1}, \ldots, 0$ and $a_{m_0+1,n,q} = \xi_q(r_n) r_n$, $r_n = (n+1)/N$, $n = 0, 1, \ldots, N-1$. h and g are the low pass and the high pass FIR filters, respectively. 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.

The average 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

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

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

4. Why does the filter bank perform better than the conventional methods?

Let us illustrate the use of the approximation and the detail coefficients. Consider a prototype object O_p and a test object O_t. Let us introduce a relative approximation and detail error given by

$$e^q_{A,m} = \sqrt{\sum_n \left(\frac{\left|a^{O_p}_{m,n,q}\right| - \left|a^{O_t}_{m,n,q}\right|}{\left|a^{O_p}_{m,n,q}\right|} \right)^2}, \qquad e^q_{D,m} = \sqrt{\sum_n \left(\frac{\left|d^{O_p}_{m,n,q}\right| - \left|d^{O_t}_{m,n,q}\right|}{\left|d^{O_p}_{m,n,q}\right|} \right)^2}.$$

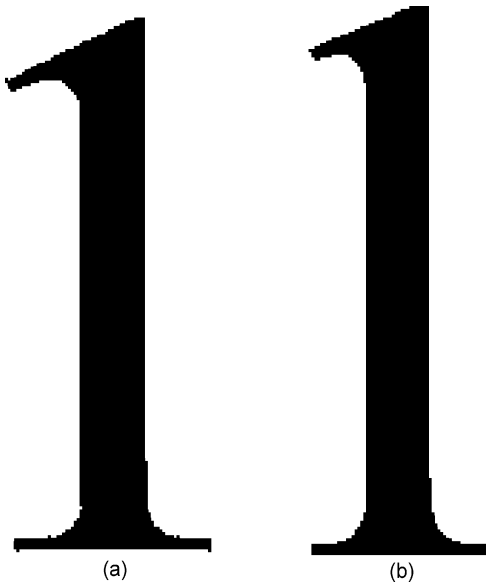

Fig. 2. The digit and letter. (a) 1 and (b) "l".

Recall that q is the index of the angular order, m corresponds to the resolution level and n to the spatial position of the wavelet. The estimates $e^q_{A,m}$ and $e^q_{D,m}$ are nothing else than norms of vectors composed from the relative errors in the approximation and detail coefficients at decomposition level m.

Our examples show that in some cases only the approximation error should be used for recognition. In some other cases it should be the detail error and, in general, the pattern recognition systems should employ combinations of the detail and approximation coefficients. Fig. 2 shows two different objects, digit 1 and letter "l". Digit 1 which has been subjected to the boundary noise 5% and 10% is depicted in Fig. 3(a) and (b). An object created from digit 1 by adding a round attachment in the middle is depicted in Fig. 4. We also analyzed nine objects obtained from digit 1 by cutting out a round hole having a varying location (see Fig. 5(a)–(i)). Finally, two different letters "O" and "Q" are depicted in Fig. 6.

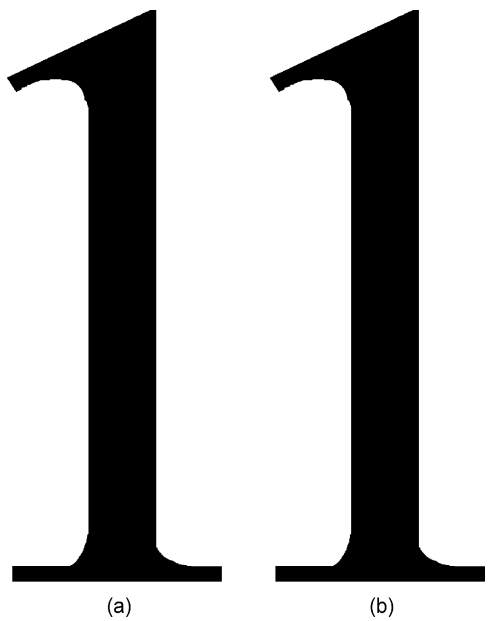

Fig. 3. The digit 1 with boundary noise. (a) 5% noise and (b) 10% noise.

Fig. 4. The digit 1 with attachment in the middle.

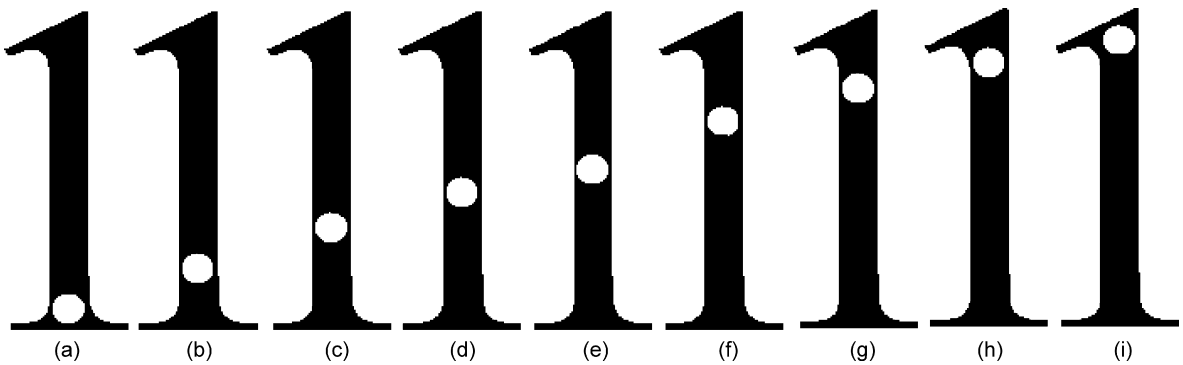

Fig. 5. The digit 1 with a hole having varying location inside the digit.

Tables 1–7 show the relative errors in the approximation and the detail coefficients corresponding to a five level multiresolution analysis by means of the B-spline wavelets, the angular order $q = 1$. Consider, now vectors e_A and e_D composed of elements $e^1_{A,m}$ and $e^1_{D,m}$. For the case of digit 1 and letter "l" the details are slightly better for every resolution level (for instance, for $m = 5$, $e_A = 22.1882$, $e_D = 25.4783$).

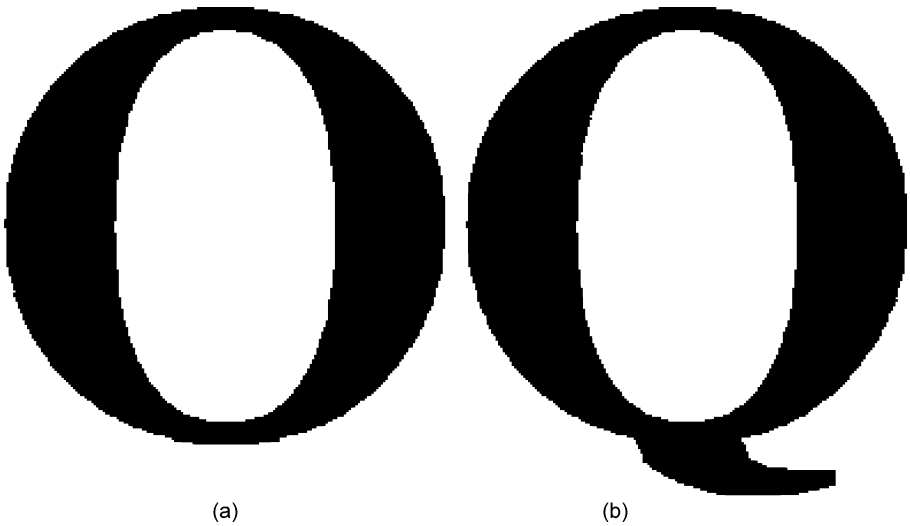

Fig. 6. The letters (a) "O" and (b) "Q".

Table 1
Relative error: 1 and "l".

e	m				
	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$
$e^1_{A,m}$	22.1882	6.9889	5.8049	2.7225	1.6057
$e^1_{D,m}$	25.4783	11.6864	5.7027	3.2314	1.7271

Table 2
Relative error: 1 with 5% boundary noise.

e	m				
	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$
$e^1_{A,m}$	6.2042	2.6995	1.7308	0.9459	0.5349
$e^1_{D,m}$	9.1066	4.3081	1.4032	0.7560	0.5998

Table 3
Relative error: 1 with 10% boundary noise.

e	m				
	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$
$e^1_{A,m}$	11.3168	6.7803	3.3334	1.5894	0.9619
$e^1_{D,m}$	18.1761	9.8956	7.5661	2.0695	2.0016

Table 4
Relative error, 1 with an attachment.

e	m				
	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$
$e^1_{A,m}$	42.3518	20.4168	11.6810	6.8006	4.6799
$e^1_{D,m}$	29.4903	4.7810	7.7067	5.8521	2.0794

Table 5
Relative error, approximations, 1 with a hole having varying location.

Position in Fig. 5	$e^1_{A,m}$					Total error
	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$	
(i)	7.8313	5.1957	3.6077	1.7937	0.5443	18.9727
(h)	9.9128	6.1723	2.9294	2.1135	1.0101	22.1381
(g)	10.4518	5.6071	3.9271	4.5559	2.3707	26.9126
(f)	52.6387	30.1209	24.7501	30.1129	15.4706	153.0932
(e)	29.5425	12.9450	10.1065	19.1939	9.3009	81.0888
(d)	11.1161	5.7834	2.4427	3.2869	1.5367	24.1658
(c)	8.4864	4.5070	1.5316	1.1410	0.8624	16.5284
(b)	8.1737	4.5867	2.2386	1.7691	0.6853	17.4534
(a)	6.2806	3.7115	2.1025	0.9009	0.3651	13.3606
Average	16.0482	8.7366	5.9596	7.2075	3.5718	

Table 6
Relative error, details, 1 with a hole having varying location.

Position in Fig. 5	$e^1_{A,m}$					Total error
	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$	
(i)	16.8718	15.6832	9.4616	5.5532	5.1424	52.7122
(h)	24.1564	20.3270	15.2146	7.1422	5.8996	72.7398
(g)	19.7997	13.4352	4.9179	5.6855	2.8146	46.6529
(f)	11.0551	4.5839	4.0826	3.1346	1.8397	24.6959
(e)	6.9707	1.9645	1.7970	1.1650	1.6512	13.5484
(d)	19.9375	10.9415	9.1181	5.7682	1.8442	47.6095
(c)	17.2088	13.3634	9.8229	6.2616	3.9144	50.5711
(b)	18.8860	11.8602	8.2078	7.2531	8.1094	54.3165
(a)	11.9447	8.0676	6.4575	4.9301	2.5286	33.9285
Average	16.3145	11.1363	7.6756	5.2104	3.7493	

However, digit 1 subjected to 5% and 10% boundary noise is characterized by $e_A = 6.2042$, $e_D = 9.1066$ and $e_A = 11.3168$, $e_D = 18.1761$. Moreover, the approximation coefficients perform better on every level (Tables 2 and 3).

Small detail errors often appear in cases when a new object is obtained from the original object by attaching an additional part or cutting holes inside the object. An object obtained from digit 1 by adding an attachment shown in Fig. 4 is characterized by $e_A = 42.3518$, $e_D = 29.4903$. The approximation error is larger for each resolution level (Table 4). An interesting experiment with a round hole having varying location inside the digit 1 is presented in Tables 5 and 6.

Tables 5 and 6 reveal that the behavior of the error associated with the approximation and detail coefficient is noticeably changing depending on the location of the hole. Suppose that we consider the new objects different from digit 1. Then the approximation coefficients perform the best when the hole is placed near the center of the object (Table 5). Interestingly enough the performance of the detail coefficients is the opposite (Table 6).

The best recognition occurs when the hole is located at the periphery of the object. This phenomenon has a simple explanation.

Since the object has virtually the shape of a rectangle with a high aspect ratio the hole may significantly contribute to the approximation error (low frequency) when it is close to the centroid. However, the hole far from the centroid changes the result of integration only for large r. Since for these r-s there exists a large area belonging to the white background, the hole affects only the high frequencies and the details become more important than the approximations at all levels.

Finally, the approximation and detail errors corresponding to two different letters "O" and "Q" are given in Table 7. The approximation is better at every resolution level ($e_A = 89.8708$, $e_D = 48.1100$).

The examples clearly demonstrate that feature selection must be performed with regard to the approximation and detail coefficients. It may consist of only the details as in [27], or only of the approximations such as in the case of the object in Figs. 4 and 6. However, usually it is a combination of approximations and details calculated for at several resolution levels. Besides, the feature selection procedure must be intelligent enough to use the multiresolution nature of the filter bank to analyze not only the frequency bands but contributions of the individual wavelets as well. As a matter of fact, this problem has not been approached in connection to wavelet moments [27] before. The next section presents one of the possible solutions.

Table 7
Relative error: "O" and "Q".

e	m				
	$m=5$	$m=4$	$m=3$	$m=2$	$m=1$
$e^1_{A,m}$	89.8708	61.7995	19.9501	21.0388	13.4470
$e^1_{D,m}$	48.1100	31.7471	15.1061	12.7774	20.1356

5. Feature selection algorithm

Selection of features is a crucial step for an object recognition system [1,30]. The aim is to generate the best combination of features that maximizes the recognition rate. We present an algorithm based on examining the bands and the features individually and in combinations. As opposed to [27] our algorithm takes a 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 to analyze combinations of the wavelets. A detailed description of the feature selection procedure is given below.

(1) Calculate angular orders by sampling 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

$$E_{ls}(q) = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{n=1}^{N}\left(\left|\xi_q(r_n)^{i,Template}\right| - \left|\xi_q(r_n)^{i,j}\right|\right)^2}{I\,J\,N}, \tag{5}$$

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. Note that the procedure is just a preliminary filtering which allows to discard easily recognizable bad choices defined by $E(q) \geq \varepsilon_{\text{pre}}$, where ε_{pre} is a prescribed threshold. The resulting set of all q' that pass the pre-filtering is fed to the next step of the procedure.

(3) Apply the QMF to q' as illustrated in Fig. 7.
The result is a tree-structured filter bank of coefficients (features).

(4) Prune the resulting filter bank. The algorithm finds the best discriminating subbands which produce well-separated classes. A symmetric version of the Kullback–Leibler distance based on the relative entropy [26,31] is used to measure the discrimination power of the subband. A good subband is the one that reduces the relative entropy [31] which for the case of two classes C_1 and C_2 is given by

$$\delta_{C_1C_2}(m,k,q) = 1/2 \sum_n \left(\gamma_{C_1}(m,k,n,q)\log\left(\frac{\gamma_{C_1}(m,k,n,q)}{\gamma_{C_2}(m,k,n,q)}\right) + \gamma_{C_2}(m,k,n,q)\log\left(\frac{\gamma_{C_2}(m,k,n,q)}{\gamma_{C_1}(m,k,n,q)}\right)\right) \tag{6}$$

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

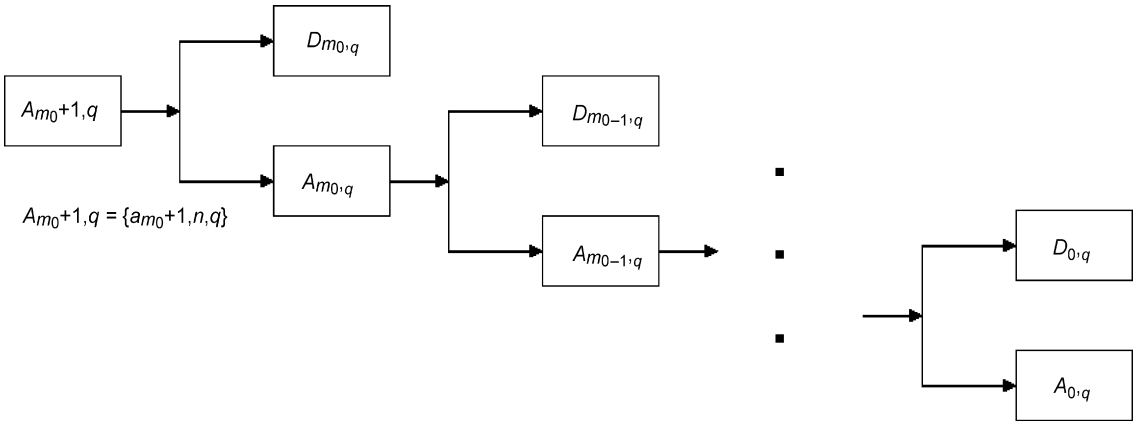

Fig. 7. The two band filter bank.

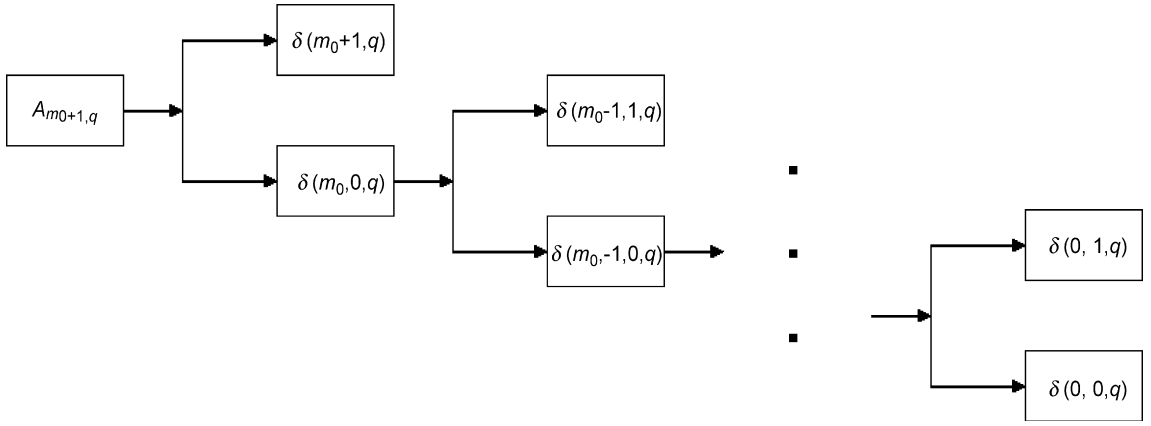

Fig. 8. The Kullback–Leibler distance (relative entropy), a two-band filter bank.

* denotes complex conjugate, m is the resolution level, $k = 0, 1$ denotes the left and right branches of the tree, n is the position index. $w_{m,n,q}$ is given by

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

In the case of I classes, $\delta(m, k, q) = \sum_{i=1}^{I-1}\sum_{j=i+1}^{I}\delta_{C_iC_j}(m, k, q)$ (Fig. 8).

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

(5) Reduce the dimension of the feature space by analyzing the features individually using a statistical testing ANOVA [14,32]. We use a one-way ANOVA with a randomized complete block design to verify the assumption $\mu_1 \neq \mu_2 \neq \cdots \mu_i \neq \cdots \neq \mu_I$, 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. 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

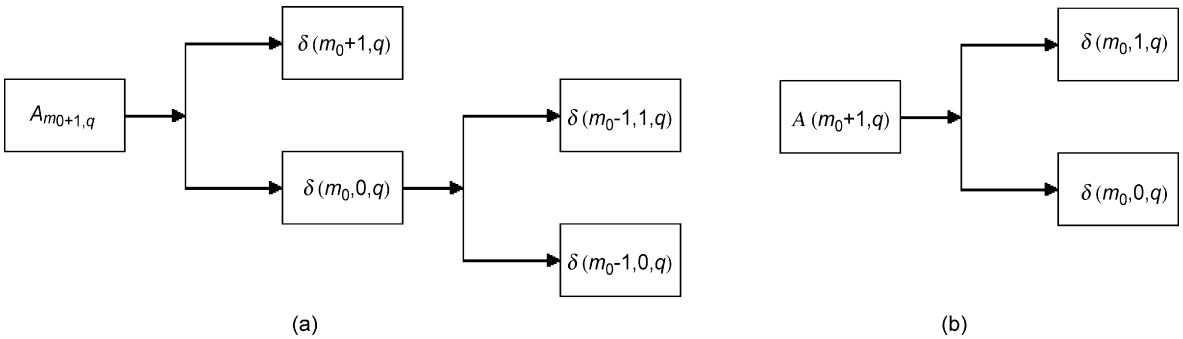

Fig. 9. Example of the pruning algorithm for two main cases: Case (a): $\delta(m_0, 0, q) > \delta(m_0 - 1, 0, q) + \delta(m_0 - 1, 1, q)$. Keep subbands $\delta(m_0 - 1, 0, q)$ and $\delta(m_0 - 1, 1, q)$; Case (b): $\delta(m_0, 0, q) \leq \delta(m_0 - 1, 0, q) + \delta(m_0 - 1, 1, q)$. Remove subbands $\delta(m_0 - 1, 0, q)$ and $\delta(m_0 - 1, 1, q)$.

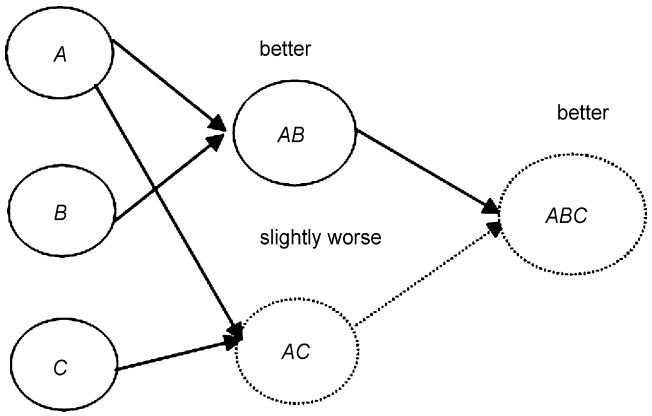

Fig. 10. Conventional Apriori algorithm (the solid line) and modified Apriori algorithm (the solid and the dashed lines), "better" means that the combination is better than the both features separately.

based on evaluating the partially supervised FCM-type cost function [2,7,9,11,23,24,28] used by analogy with the measure of confidence in the conventional Apriori algorithm [10].

The cost function is given by

$$f_C(X) = J(X)\ \log(N_{miss} + \vartheta), \tag{7}$$

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 standartized 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

$$f_C(A \cup B) \leq f_C(A) \quad \text{and} \quad f_C(A \cup B) \leq f_C(B).$$

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

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

where defines Δ the allowable interval of confidence (Fig. 10).

The confidence in a combination of n features is evaluated using the same principle. It should be noted that the features must be normalized and standardized prior to the selection. The best combination is the k-itemset which minimizes the cost function above. It is also possible to select several itemsets with the smallest values of the cost function. Finally, the MAA 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'.

6. Why does the feature selection perform better?

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

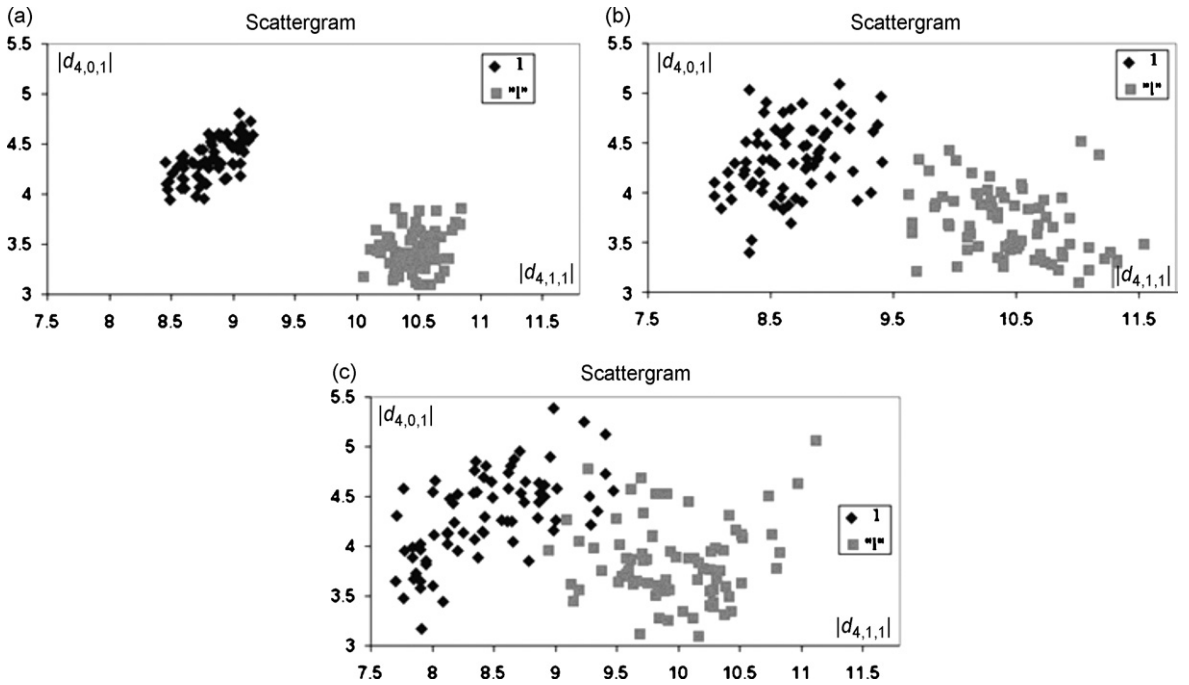

Fig. 11. Noise sensitive features extracted from the digit 1 and letter "l". (a) Low level noise, (b) medium level noise, and (c) high level noise.

6.1. Example 1: noise sensitive features

Noise sensitive features $|d_{4,0,1}|$ and $|d_{4,1,1}|$ extracted 1 and "1" are shown in Fig. 11 The high level noise makes the objects inseparable. Consequently, $|d_{4,0,1}|$ is discarded at the ANOVA stage. $|d_{4,1,1}|$ has a better separability, however, this feature will be discarded as well at the Apriori selection stage.

6.2. Example 2: pruning the filter bank tree

Fig. 12 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 $m = 1$ as well as level $m = 3$ have been eliminated since $(\delta(a, 2) = 0.082) < (\delta(a, 1) + \delta(d, 1) = 0.098)$ and $(\delta(a, 4) = 0.394) < (\delta(a, 3) + \delta(d, 3) = 0.423)$.

Note that although level $m = 3$ has been excluded, it does not mean that all the subsequent levels are not appropriate. For example, in this case, level $m = 2$ has also been selected (Fig. 13).

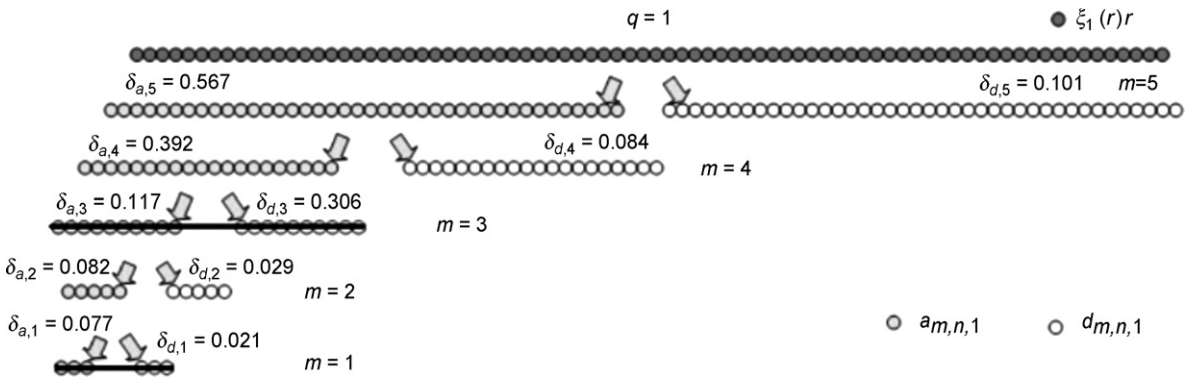

Fig. 12. Pruning subbands at resolution level $m = 1$ and $m = 3$ by using the Kullback–Leibler distance.

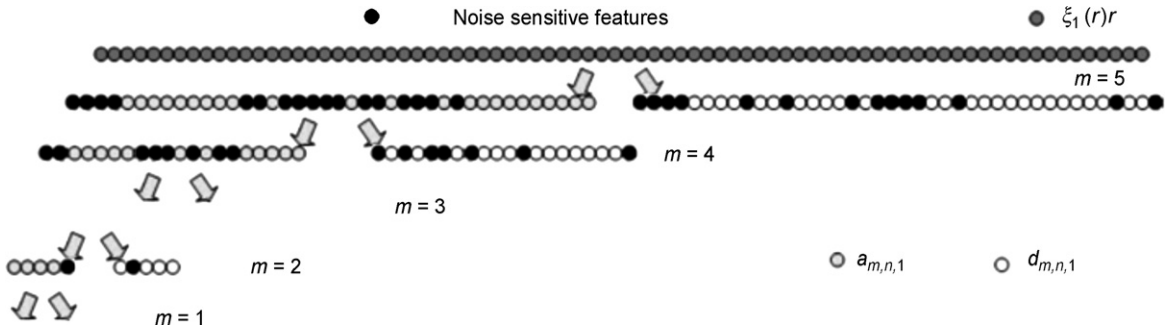

Fig. 13. Discarding noise sensitive features (for every particular subband) by using ANOVA.

6.3. Example 3: the modified Apriori algorithm

Our 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 $m = 4$ consisting of 40 coefficients (20 details and 20 approximations). The total number of combinations is approximately 1.1×10^{11}. Assuming that one combination requires only one floating point operations we obtain approximately 1000 Gflops.

L_1	L_2	$fc,_2$	L_3	$fc,_3$	L_4	$fc,_4$
$\{a_2\}$	$\{a_3, a_5\}$	31.8241	$\{a_3, a_5, d_9\}$	32.4627	$\{a_{15}, a_{16}, d_{14}, d_{17}\}$	**16.8072**
$\{a_3\}$	$\{a_3, a_6\}$	34.0666	$\{a_{15}, a_{16}, d_8\}$	19.7274		
$\{a_4\}$	$\{a_3, d_9\}$	32.8229	$\{a_{15}, a_{16}, d_{14}\}$	18.1456		
$\{a_5\}$	$\{a_5, a_6\}$	33.1313	$\{a_{15}, a_{16}, d_{17}\}$	17.9963		
$\{a_6\}$	$\{a_5, d_9\}$	31.3869	$\{a_{15}, d_8, d_{14}\}$	20.5647		
$\{a_{10}\}$	$\{a_6, d_9\}$	34.9852	$\{a_{15}, d_8, d_{17}\}$	20.1226		
$\{a_{12}\}$	$\{a_{15}, a_{16}\}$	19.0877	$\{a_{15}, d_{14}, d_{17}\}$	**15.9086**		
$\{a_{15}\}$	$\{a_{15}, d_8\}$	20.3031	$\{a_{16}, d_8, d_{14}\}$	19.1925		
$\{a_{16}\}$	$\{a_{15}, d_{14}\}$	**16.8451**	$\{a_{16}, d_8, d_{17}\}$	19.9795		
$\{a_{17}\}$	$\{a_{15}, d_{17}\}$	**16.4754**	$\{a_{16}, d_{14}, d_{17}\}$	17.8878		
$\{a_{18}\}$	$\{a_{16}, d_8\}$	20.4865	$\{d_8, d_{14}, d_{17}\}$	19.1137		
$\{a_{19}\}$	$\{a_{16}, d_{14}\}$	19.8297				
$\{d_1\}$	$\{a_{16}, d_{17}\}$	18.8300				
$\{d_3\}$	$\{d_8, d_{14}\}$	20.3291				
$\{d_6\}$	$\{d_8, d_{17}\}$	20.1562				
$\{d_8\}$	$\{d_{14}, d_{17}\}$	**16.8140**				
$\{d_9\}$						
$\{d_{10}\}$						
$\{d_{12}\}$						
$\{d_{13}\}$						
$\{d_{14}\}$						
$\{d_{15}\}$						
$\{d_{16}\}$						
$\{d_{17}\}$						
$\{d_{18}\}$						

Fig. 14. The modified Apriori algorithm applied to a filter bank of rotationally invariant moments.

Fig. 14 illustrates the MAA applied to this case. The first itemset L_1 is the output of ANOVA. It consists of 12 approximation and 13 detail coefficients. The next level L_2 contains only 14 items, L_3—11 items and L_4 has only 1 item. The number of the floating point operations required at the first step is equal to 0.3 Gflops, at the second step it is 0.0006 Gflops, finally the third step consumes only 0.000002 Gflops. Therefore, the time required for calculations has been reduced by a factor of 3000.

Suppose that the final set of the best combinations of features is selected using the condition $f_C \leq 17$. It includes then $\{a_{15}, d_{14}\}$, $\{a_{15}, d_{17}\}$, $\{d_{14}, d_{17}\}$, $\{a_{15}, d_{14}, d_{17}\}$ and $\{a_{15}, a_{16}, d_{14}, d_{17}\}$. The best combination is $\{a_{15}, d_{14}, d_{17}\}$. 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 $f_C \leq 17$ will not be produced.

However, the MAA characterized by $\Delta = 2$ will generate $\{a_{15}, a_{16}, d_{14}, d_{17}\}$. Note that we the case of $20 + 20$ coefficients above was considered for convenience. 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.

7. Experimental results

We evaluate the performance of the proposed algorithm by three datasets. The first dataset consists of 48,400 noisy images [34] based on 20 basic aircraft silhouettes: 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, and Yak-36 (see Fig. 15). Each silhouette produces 1600 training images and 820 testing images. The second dataset based on an online database NIST [8], consists of machine-printed characters, namely, 11,000 upper case English letters (bold, courier). We use 7000 letters for training and 4000 for testing. The last dataset consists of 14,000 gray level images based on 10 Thai musical instruments [4]: SAW DUANG, SAW OU, SAW SAM SAI, SALOR, SUENG, JAKAE, PEE CHAWA, PEE NOKE, RANAD TUM, and RANAD AKE (see Fig. 16). Each instrument produces 950 training images and 450 testing images. All of the datasets are degraded by an impulse noise varying from 0% to 8% and a transformation noise. In the case of the NIST data we also consider an interesting effect of the boundary noise appearing after separation of touching letters by means of dilation.

We discuss experiments with the B-spline wavelets, however, the orthogonal wavelets such as the Daubechies wavelets 2, 4 and 6 and the Coiflet wavelets [3,18] were tested as well. Although the orthogonal wavelets easily allow to reconstruct the image, the biorthogonal B-splines with underlying symmetric FIR filters 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 the cost function given in Section 5.

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 if the features were selected individually based on the between-to within-class variance ratio [27].

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 8. Table 8 includes degradation by all types of noise: rotation, translation, scaling and random valued impulse noise which η denotes the impulse noise percentage.

Table 8 shows the 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 improves by 9%.

Tables 9–14 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 every case algorithm always outperforms Shen-I-V, Shen-FCM-M, as well as Zernike-I-V and Fourier–Mellin-I-V.

The efficiency of the 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-I-V in the case of the aircraft silhouettes degraded by 6–8% impulse noise and the rotation noise (Table 10). The rotation and segmentation noise affects the NIST characters and the Thai musical instruments

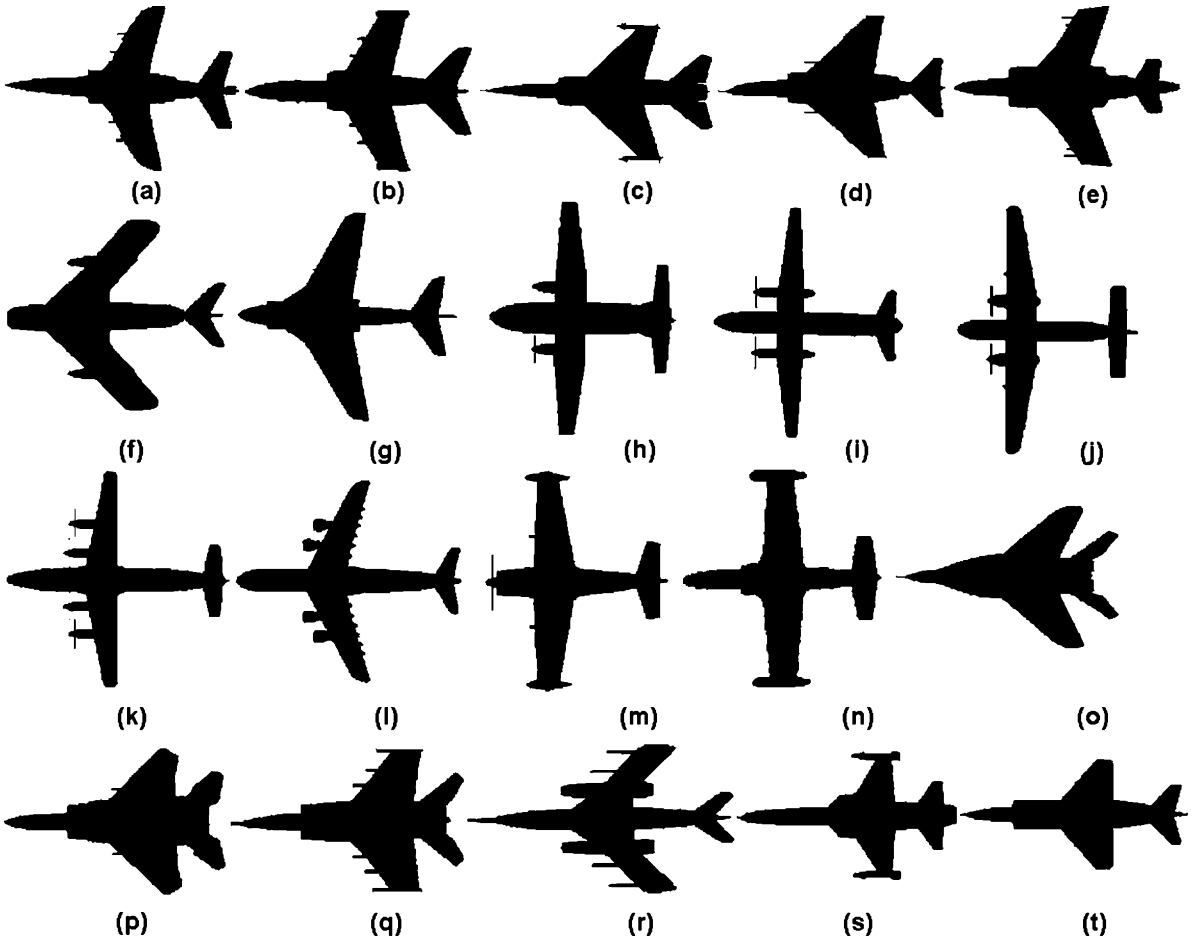

Fig. 15. Twenty Silhouettes of aircraft. (a) Alpha Jet, (b) Am-X, (c) Mirage F1, (d) F-4 Phantom, (e) F-15 Eagle, (f) MiG-17, (g) A-6 Intruder, (h) Aviocar C-212, (i) An-32 Cline, (j) F-5 Freedom, (k) An-12 Cub, (l) Hunter, (m) Brewer, (n) Jastreb, (o) MiG-29, (p) Buccaneer, (q) MiG-25, (r) Mirage III, (s) F-18 Hornet, and (t) Yak-36.

(Tables 11–14) more significant than the aircraft silhouettes since the centroids of the characters and the instruments often lie outside the object body.

Consequently, the centroids are much more sensitive to the noise. Observe the most interesting cases. In case of the NIST characters (see Table 12, 4.5–6% noise) and a combination of the impulse noise and the transformation noise

Table 8
Average classification rates.

Algorithms	Classification rate		
	Aircraft silhouettes $0 \leq \eta < 8\%$	Upper case NIST character $0 \leq \eta < 6\%$	Thai musical instruments $0 \leq \eta < 4.5\%$
FB-P-AN-MAA-FCM-P	95.38	95.46	93.07
FB-P-AN-MAA-FCM-M	91.19	92.76	89.78
FB-P-AN-MAA-FCM-E	88.05	89.73	84.66
Shen-FCM-M	88.51	90.94	86.58
Shen-FCM-E	85.17	89.10	83.86
Shen-I-V	82.94	85.94	81.01
Zernike-I-V	82.19	85.03	80.51
Fourier–Mellin-I-V	77.61	79.31	74.16

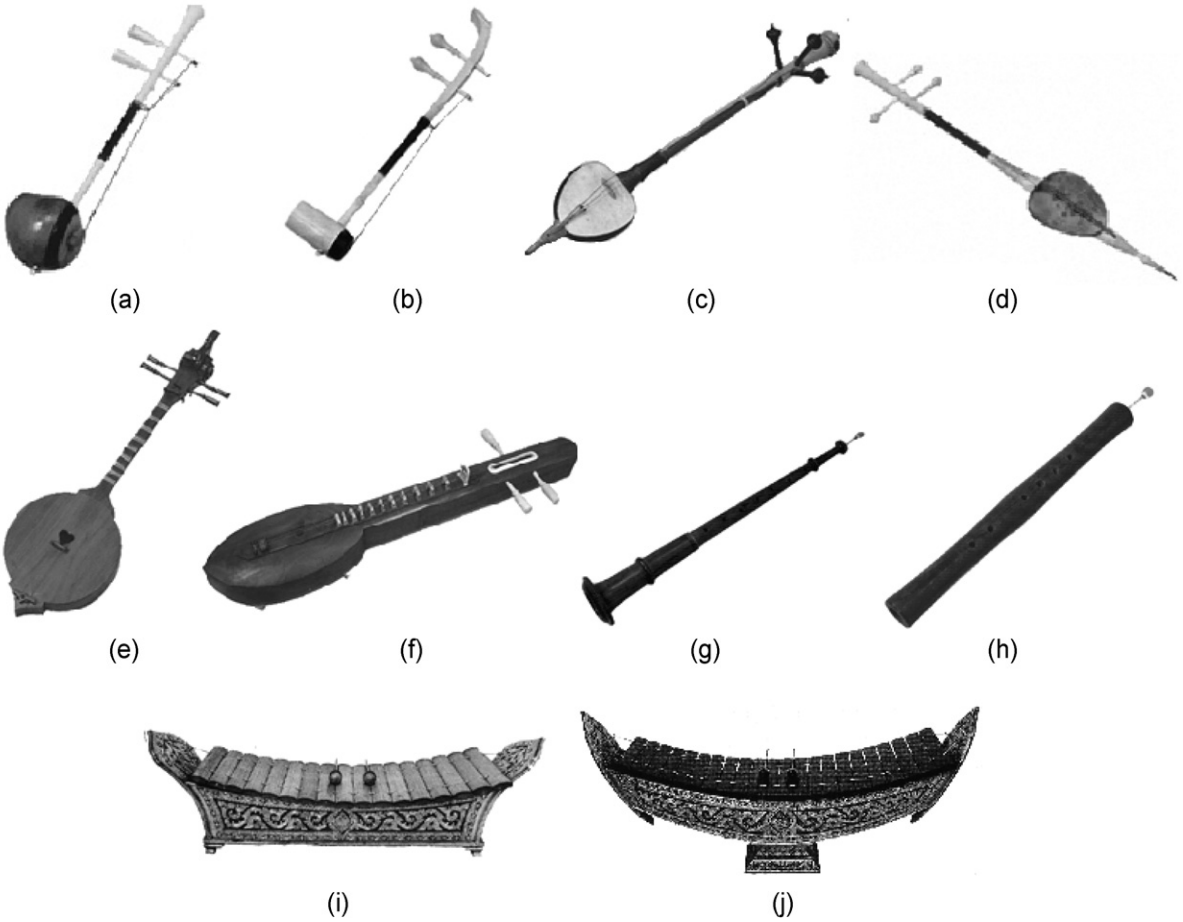

Fig. 16. Gray level images of Thai musical instruments. (a) "SAW DUANG" (Fiddle), (b) "SAW OU" (Fiddle), (c) "SALOR" (Lute), (d) "SAW SAM SAI" (Fiddle), (e) "SUENG" (Lute), (f) "JAKAE" (Lute), (g) "PEE CHAWA" (Pipe), (h) "PEE NOKE" (Pipe), (i) "RANAD TUM" (gamelan), and (j) "RANAD AKE" (gamelan).

Table 9
Aircraft images, impulse noise.

Algorithms	Noise $0 \leq \eta < 2$	Noise $2 \leq \eta < 4$	Noise $4 \leq \eta < 6$	Noise $6 \leq \eta < 8$
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 10
Aircraft images, impulse noise combined with rotation and scaling.

Algorithms	Noise $0 \leq \eta < 2$	Noise $2 \leq \eta < 4$	Noise $4 \leq \eta < 6$	Noise $6 \leq \eta < 8$
FB-P-AN-MAA-FCM-P	98.45	94.66	90.28	78.53
QMF-PRUNING-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	82.30	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 11
The NIST characters, impulse noise and segmentation noise.

Algorithms	Noise $0 \leq \eta < 1.5$	Noise $1.5 \leq \eta < 3$	Noise $3 \leq \eta < 4.5$	Noise $4.5 \leq \eta < 6$
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	96.90	87.73	78.03	58.24
Fourier–Mellin-I-V	94.27	81.82	70.12	49.61

Table 12
The NIST characters, impulse noise and transformation noise.

Algorithms	Noise $0 \leq \eta < 1.5$	Noise $1.5 \leq \eta < 3$	Noise $3 \leq \eta < 4.5$	Noise $4.5 \leq \eta < 6$
FB-P-AN-MAA-FCM-P	98.85	94.40	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	89.60	77.08	61.12	39.90

Table 13
Gray level images of Thai musical instruments, random-valued impulse noise.

Algorithms	Noise $0 \leq \eta < 1.5$	Noise $1.5 \leq \eta < 3$	Noise $3 \leq \eta < 4.5$
FB-P-AN-MAA-FCM-P	97.29	92.54	85.13
FB-P-AN-MAA-FCM-M	95.02	87.38	77.98
Shen-FCM-M	93.81	86.53	75.02
Shen-I-V	92.04	82.18	71.11
Zernike-I-V	91.74	81.20	70.96
Fourier–Mellin-I-V	89.02	77.49	67.82

Table 14
Gray level images of Thai musical instruments, random-valued impulse noise and transformation noise.

Algorithms	Noise $0 \leq \eta < 1.5$	Noise $1.5 \leq \eta < 3$	Noise $3 \leq \eta < 4.5$
FB-P-AN-MAA-FCM-P	96.53	89.14	78.44
FB-P-AN-MAA-FCM-M	92.71	80.07	67.08
Shen-FCM-M	91.04	76.21	62.24
Shen-I-V	89.82	74.52	58.51
Zernike-I-V	89.88	70.19	56.96
Fourier–Mellin-I-V	86.17	65.32	49.13

Shen-I-V, Zernike-I-V and Fourier–Mellin-I-V display 50.93%, 46.31% and 39.90% recognition rate, respectively whereas the proposed filter bank invariants 80.51%. In case of the Thai musical instruments and 3–4.5% random-valued impulse noise combined with the transformation noise the Shen-I-V, Zernike-I-V and Fourier–Mellin-I-V display 58.51%, 56.96% and 49.13% recognition rate, respectively whereas the proposed filter bank invariants show 78.44% recognition rate.

8. 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 postdoctoral scholarship (grant MRG4980082) awarded to Dr. A. Rodtook by the Thailand Research Fund (TRF) and the Commission on Higher Education (CHE).

References

[1] S. Abe, Optimizing Features, Springer, 1984.
[2] J. Bezdek, Pattern Recognition with Objective Function Algorithms, Plenum, 1981.
[3] C. Burrus, R. Gopinath, H. Guo, Introduction to Wavelets and Wavelets Transforms, Prentice Hall, NJ, USA, 1998.
[4] P. Charlerycharoen, Visual Database of Thai Musical Instruments, Wattanapanit, Bangkok, TH, 2000.
[5] J. Flusser, On the independence of rotation moment invariants, Pattern Recognition 33 (2000) 1405–1410.
[6] J. Flusser, On the inverse problem of rotation moment invariants, Pattern Recognition 35 (2002) 3015–3017.
[7] G. Frosini, B. Lazzerini, F. Marcelloni, A modified fuzzy C-means algorithm for feature selection, in: Proceeding of the 19th International Conference of the North American Fuzzy Information Processing Society: NAFIPS-2000, Atlanta, GA, USA, 2000, pp. 148–152.
[8] M. Garris, NIST Special Database 8: NIST Machine-Print Database of Gray Scale and Binary Images, National Institute of Standards and Technology, Gaithersburg, MD, USA. http://www.nist.gov/srd/nistsd8.htm.
[9] D. Gustafson, W. Kessel, Fuzzy clustering with a fuzzy covariance matrix, in: Proceeding of the International Conference on Decision and Control: IEEE CDC-78, 1978, pp. 761–766.
[10] J. Han, M. Kamber, Data Mining, chap. Mining Association Rules in Large Databases, Morgan Kaufmann, San Francisco, USA, 1984, pp. 225–269.
[11] R.J. Hathaway, J.C. Bezdek, Recent convergence results for the fuzzy c-mean clustering algorithms, Journal of Classification 5 (1988) 237–247.
[12] M.K. Hu, K.E. Barner, Visual pattern recognition by moment invariants, IRE Transaction Information Theory 8 (1962) 179–187.
[14] R.A. Johnson, Probability and Statistics for Engineers, chap. Analysis of Variance, 6th ed., Prentice Hall, NJ, USA, 2000, pp. 407–414.
[15] C. Kan, M.D. Srinath, Invariant character recognition with zernike and orthogonal Fourier–Mellin moments, Pattern Recognition 35 (2002) 143–154.
[16] S.X. Liao, M. Pawlak, On the accuracy of zernike moments for image analysis, IEEE Transaction on Pattern Analysis and Machine Intelligence 20 (12) (1998) 1358–1364.
[17] S. Maitra, Moment invariants, Proceedings of the IEEE Publication 67 (4) (1979) 697–699.
[18] S.G. Mallat, Multifrequency channel decompositions of image and wavelet models, IEEE Transaction on Image Processing 37 (12) (1989) 2091–2110.
[20] R. Mukundan, S.H. Ong, P.A. Lee, Image analysis by tchebichef moments, IEEE Transaction on Image Processing 10 (9) (2001) 1357–1364.
[21] R. Mukundan, K.R. Ramakrishnan, Fast computation of legendre and zernike moments, Pattern Recognition 28 (9) (1995) 1433–1442.
[22] Y.R. Paramesran, S.H. Ong, Image analysis by krawtchouk moments, IEEE Transaction on Image Processing 12 (11) (2003) 1367–1377.
[23] W. Pedrycz, Algorithms of fuzzy clustering with partial supervision, Pattern Recognition Letter 3 (1985) 13–20.
[24] W. Pedrycz, J. Waletzky, Fuzzy clustering with partial supervision, IEEE Transaction on Systems, Man and Cybernetics 27 (5) (1997) 787–795.
[25] S. Rodtook, S.S. Makhanov, Numerical experiments on the accuracy of rotation moment invariants, Image and Vision Computing 23 (2005) 577–586.
[26] N. Saito, R.R. Coifman, Local discriminant bases, in: Proceeding of SPIE 2303, Wavelet Applications in Signal and Image Processing, San Diego, USA, 1994, pp. 2–14.
[27] D. Shen, H.H. Ip, Discriminative wavelet shape descriptors for recognition of 2-d patterns, Pattern Recognition 32 (1999) 151–165.
[28] H.J. Sun, S.R. Wang, Z. Mei, A fuzzy clustering based algorithm for feature selection, in: Proceeding of the 1st International Conference Machine Learning and Cybernetics, Beijing, China, 1993–1997.
[29] C.H. Teh, R.T. Chin, On image analysis by the methods of moments, IEEE Transaction on Pattern Analysis and Machine Intelligence 10 (4) (1988) 496–512.
[30] S. Theodoridis, K. Koutroumbas, Pattern Recognition, chap. Feature Selection, Academic Press, London, UK, 1999, pp. 139–179.
[31] M. Thuillard, Wavelets in Soft Computing Series in Robotics and Intelligent Systems, chap. Spline-Based Wavelets Approximation and Compression Algorithms, World Science, NJ, USA, 2001, pp. 33–56.
[32] D. Tipyota, Probability and Statistics, chap. Analysis of Variance, CUprint Press, Bangkok, TH, 2002, pp. 329–333.
[33] M. Unser, A. Aldroubi, Wavelets-A Tutorial in Theory and Application, chap. Polynomial Splines and Wavelets—A Signal Processing Perspective, Academic Press, Boston, USA, 1992, pp. 191–122.
[34] J.A. Wickham, Visual Aircraft Recognition, Headquarters Department of the US Army, Washington, DC, USA, 1986.