Pattern Classification in a Noisy Environment

M. Mladenov

Key Words: *Classification; noisy environment; class area approximation; radial basis elements.*

Abstract*. This paper is focused on the classification of multidimensional patterns in classes located in a noisy environment using approximation of the class areas through radial basis elements. Different variants of class area approximation are proposed and investigated. An approach for classification in overlapping classes is discussed. It is based on the class areas approximation taking into account the class potentials. A variant of classifier validation, which uses noisy vectors in validating sets, is investigated. The classifiers are trained and tested with simulated data and experimental data, related to the recognition of color characteristics of grain sample elements. The results obtained by classifiers proposed and classifier with standard radial basis functions and Bayesian classifier are compared.*

1. Introduction

When we solve different classification tasks, the input vectors are grouped into classes with definite class centers, which correspond to the class prototypes. The dispersion of classes' vectors is due to two main factors: the deviation of the actual values of the object characteristics and the measurement noise. These two quantities are independent random variables in many application tasks. Usually the vector dispersion has Gaussian distribution.

The class areas in which the input vectors have to be distributed have normally limited dimensions and specific shape. In many cases the shape of the class boundaries is conditioned by the task context. Assuming that the input vectors have Gaussian distribution and vector component variances are equal along the directions of the feature space, the class areas are round shaped. This condition is not fulfilled in many classification tasks. Often the variances of the input vector components are sufficiently different. This means that the shape of the class areas can be sufficiently different from sphere. In classification task related to the quality assessment of different products, the shape of the class areas is often similar to a prolonged ellipsoid or a parallelepiped and the areas have different orientation in the feature space. Furthermore, part of the class areas can have close, contact or overlapping boundaries.

Classification problems, where the shape and the size of the classes have to be limited, are often met. Quality assessment of different products (industrial ware, agricultural products, food products, etc.) could be given as an example [1,12,13]. The quality of such products is assessed on the basis of different quality features. The values of these features form multidimensional object description. Because of the quality features vary in fixed ranges for every quality group, the class areas are limited and they have usually near located or contacted boundaries. The following problem arises when we solve such classification tasks [13]. If the class areas, which correspond to the quality groups (ω_1, ω_2 ... ω_m), are presented in the feature space (*figure 1*), that class areas are usually enveloped by an area from objects vectors, which are out of the quality groups (ω_s) (waste, sittings, impurities, etc.).

While we can obtain data about class centers and class variances for the first group of classes (from training sets), that is not possible to be made for the second group of objects. It can be supposed, that the area of these objects can fully envelope the areas of the first group of classes (*figure 1a*). Furthermore, if there are too big deviations of the actual values of the object characteristics and an intensive measurement noise, the class areas can be overlapped (*figure 1b*).

In confirmation of the discussed task actuality the following example will be presented. Within the frames of the INTECHN project (Development of intelligent technologies for assessment of quality and safety of food agricultural products), funded by Bulgarian National Science Fund, new technologies and tools for quality assessment of cereals, fruits, vegetables, milk, dairy products, meat, meat products and eggs are developed. The recognition of objects from different groups is made on the bases of data obtained by a computer vision system and a spectrophotometer. For example, when we assess the quality of a maize grain sample using computer vision system, we can

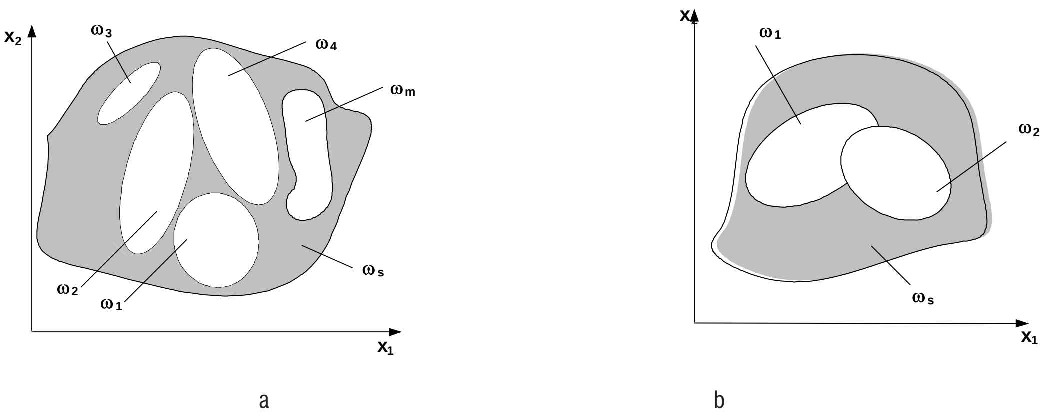

Figure 1. Class geometrical representation: a) ω_1, ω_2 ... ω_m- classes with compact areas surrounded by noisy class ω_s; b) overlapping classes surrounded by noisy class ω_s

extract information about object visible features like shape, color characteristics, dimensions, surface texture, etc. The grain sample elements can be divided in 8 basic classes (1cc - grains with inherent for the variety color, back side, 2cc - grains with inherent for the variety color, germ side, 3cc - heat-damaged grains, burned grains, 4cc - green grains, 5cc - mouldy grains, 6cc - bunt, smutty grains, 7cc - infected (with *Fusarium*) grains and 8cc - sprouted grains) on the basis of the color characteristics. These classes correspond to the color zone combinations, which are typical for different grains. There is an additional class (9cc), that corresponds to the non grain impurities. We can define comparatively compact classes for the first 8 groups. The class centers and class deviations can be calculated using the respective training sets. The investigation carried out shows that the class shapes are similar to an ellipsoid or a parallelepiped and the class axes have different orientation in the feature space. It is impossible to define a compact class for the non grain impurities because of the fact that the color and shape features of elements of this class could be sufficiently different in each subsequent grain sample.

A situation presented in *figure 1a* is arisen. In this case the classification task can be interpreted as a task for class areas approximation, when the classes are situated in a noisy environment. Furthermore, the class areas are overlapped for some of the classes *(figure1b)*. Very often correct information about class priory probabilities is missing. That makes the classification problem more complex. If we use a classifier, which requires the class priory probabilities to be known (for example Bayesian classifier [7,8]), the training procedure has to be implemented using the priory probabilities obtained from the number of elements in the training sets. When we assess quality of an unknown sample, the ratio of the number of elements in different classes can be sufficiently different from this ratio in the training sets. The classifier decision can be sufficiently different from the optimal decision under these circumstances. In this case the classification task is reduced to a task for approximation of the overlapping class areas, when the classes are situated in a noisy environment and correct data about class priory probabilities is missing.

The main goal of this work is to propose approaches and tools for effective approximation of class areas located in a noisy environment, using Radial Basis Elements (RBEs). They have to assure a correct classification of multidimensional patterns under following preconditions:

- The classes consist of Gaussian distributed vectors and they have sufficiently different dimensions along the directions of the feature space. The classes have near located class boundaries and they are situated in noisy environment;

- Class areas are overlapping and correct data about class priory probabilities is missing.

2. Classification Approach and Classifiers

2.1. Classification Approach

Having in mind the problems mentioned above, it is very important to choose an appropriate classification strategy. The use of the popular strategies, like Discriminant analysis, Cluster analysis [9], Support Vector Machines (SVM) [3,6,11], K-Nearest Neighbours (KNN) [18] and some others, which build boundaries between the class areas, is not a good choice under discussed circumstances. It would be more efficient the $\omega_1, \omega_2 \dots \omega_m$ class areas to be approximated by the classifier and all vectors, that do not get into these limited areas to be associated with the class ω_s.

Classifiers based on Radial Basis Functions [2,14,19] are appropriate in terms of the simplicity of the classification procedure and accuracy of the class areas approximation under the specific circumstances. Their application is determined by the simplicity of the classification procedure and the accuracy of the class area approximation. Furthermore, if we set an appropriate value of the RBE bias and a minimal threshold Δ of its output, it becomes clear what part of the input vectors will be included within the class boundary and it is easy to change the dimensions of the particular class area. Some of the most effective classifiers which have the ability for class area approximation are based on the Radial Basis Functions. Neural Networks (NNs), like Probabilistic Neural Network (PNN) [5,18,21] and Radial Basis Function Network (RBFN) [4,17,22], which are widely applied in similar classification tasks, include a layer with Radial Basis Elements.

2.2. Classification into Non Round/spherical Shaped Classes, which Have Sufficiently Different Dimensions in the Directions of the Feature Space and Near Located class Boundaries

Within the frames of this task we will assume, that the classes are ellipse/ellipsoid shaped. The task will be discussed in two variants - when the classes' axes are parallel to the axes of the feature space and when the class axes have a random orientation in this space.

Different variants for class area approximation using RBEs are analyzed in this investigation.

Variant 1 (CSRBE). Only one RBE is used for approximation of each class area [13]. The RBEs centers correspond to the class average values obtained from the class training sets. The class area dimensions are set through appropriate values of RBE biases and appropriate minimal threshold Δ of the RBEs outputs (*figure 2*).

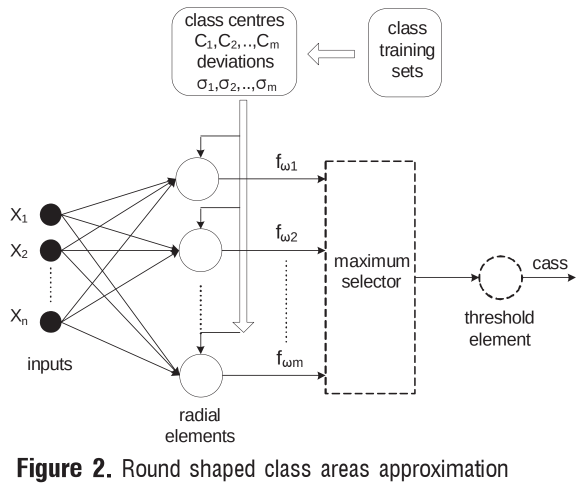

Figure 2. Round shaped class areas approximation using standard RBEs (CSRBE)

In the case, when one standard RBE is used for each class area approximation, the class boundaries are round shaped. This corresponds to the case, when the standard deviations of the input vector components are equal in all directions of the feature space. If this precondition is not carried out, then the class area shape can be sufficiently different from a sphere.

The following denotations are used: $f_{\omega i}$ is the output of i-th RBE, which corresponds to the class ω_i; Δ is threshold value, which limits class area dimensions.

The bias value of the i-th RBE is determined by the equation:

$$(1) \quad b_i = \frac{0.833}{(k \cdot \sigma_{\omega_i})}$$

where σ_{ω_i} is the standard deviation of vectors of the class ω_i; k is a parameter, which determines the dimension of the class boundary surface.

When the classes have non spherical shape, sufficiently different dimensions along the direction of the coordinate axes and near located class centers, we can expect that the CSRBE will give an incorrect classification for a part of the input vectors. The condition for correct classification of an input vector into class ω_i can be presented by the following inequalitie

(2) $f_{\omega i} > f_{\omega j}$ $(j = ... m, j \neq i)$ and

(3) $f_{\omega i} \geq \Delta$

The CSRBE can be considered as a reference classifier, because its architecture includes standard RBEs and it realizes the classification approach described in section 2.1.

Variant 2 (CDRBE1). Often the classes' shapes are sufficiently different from circle/ sphere. This case is typical for different applied tasks. If we approximate the real class areas using the CSRBE, we create preconditions for incorrect classification of a part of input vectors. A question arises: Is there a possibility the CSRBE architecture to be modified in such a way, that it can define non spherical shaped classes, for instance ellipse/ellipsoid shaped classes? A decision of this task gives the architecture of a modified classifier with decomposing RBEs (CDRBE1), presented in *figure 3*.

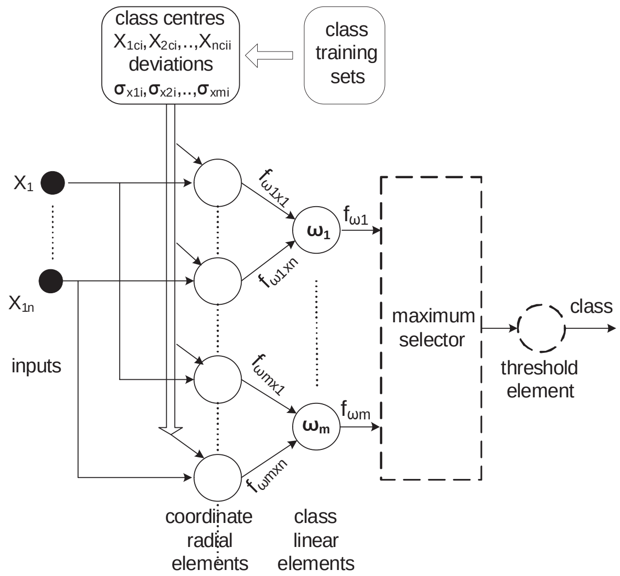

Figure 3. Classifier with decomposing RBEs (CDRBE1)

The CDRBE1 architecture consists of two layers. The first layer includes $n \times m$ radial neurons, which are distributed into m sub layers (m is the number of the classes).). The number of RBEs in each sub layer is equal to n (n is the input vector dimensionality). Each RBE has one input, connected with the respective input vector coordinate. The weights of neurons in each sub layer are equal to the coordinates of the class center. The biases of the radial neurons correspond to the real class dimensions (standard deviations $\sigma_{x\omega i}$, $\sigma_{y\omega i}$ and $\sigma_{z\omega i}$ of the vectors from the training set).

The radial layer architecture proposed gives a possibility to form class areas, whose dimensions are different along the different coordinate axes and correspond to the dimension of the real class areas. Really the CDRBE1 decomposes the multidimensional input vectors along the coordinate axes and inserts a separate radial neuron for each coordinate axis. This gives a possibility for independent adjustment of the radial neurons in each sub layer. In this way we can define different dimension of class areas along the different coordinate axes.

The second CDRBE1 layer consists of m radial neurons. Each of them has n inputs, connected to the outputs of the respective radial neuron in the first layer. The neurons outputs present the weighted distance of the input vector to the centers of non spherical classes. To determine what class the input vector belongs to, the output $f_{\omega i}$ with maximum value is chosen. This value has to exceed the threshold Δ (the last element of the classifier architecture). Otherwise the input vector is accepted as a noise. The condition for correct classification can be presented by the following inequalities:

(4) $\| f_{\omega i} - 1 \| > \| f_{\omega j} - 1 \|$ $(j = 1.... m, j \neq i)$ and

(5) $\| f_{\omega i} - 1 \| \geq \Delta$

where $f_{\omega i} = [f_{\omega i1}, f_{\omega i2}, ..., f_{\omega in}]^T$ is a vector, whose components are the outputs of the radial neurons of the i-th sub layer of the CDRBE1, corresponding to the class ω_i.

The CDRBE1 can effectively approximate ellipse/ellipsoid shaped classes, whose axes are parallel to the coordinate axes.

Variant 3 (CDRBE2). For avoiding the problem, concerning the approximation of classes, whose axes are not parallel to the coordinate axes, it is possible the following approach to be used. A local coordinate system ($x_{\omega i}$,$y_{\omega i}$,$z_{\omega i}$, for 3-dimensional vectors), whose axes coincide with the class ω_i axes of inertia, is constructed for each of the classes. The input vector coordinates are recalculated in each of the local coordinate systems. Actually through this transformation the axes of each of the classes become parallel to the axes of the input vector coordinate system. This reduces the task, when the class axes have a random orientation in the input vector space to a task, when the axes of all classes are parallel to the feature space axes. The CDRBE1 gives a correct approximation of the classes under these preconditions.

The class dimensions regarding the local coordinate systems can be determined in the following way. The covariance matrix $COV_{\omega i}$, which corresponds to the class ω_i, can be calculated on the bases of the respective training set. The diagonal term of the $COV_{\omega i}$ consists of the variances $\sigma^2_{x\omega i}$, $\sigma^2_{y\omega i}$ and $\sigma^2_{z\omega i}$. It is possibly to recalculate these variances in the local coordinate system $x_{\omega il}y_{\omega il}z_{\omega il}$ (the recalculated variances are denoted with $\sigma^2_{x\omega il}$, $\sigma^2_{y\omega il}$ and $\sigma^2_{z\omega il}$) using the following transformation (eigenvalue equation):

(6) $COV_{\omega i}.V_{\omega i} = V_{\omega i}.D_{\omega i}$,

where $D_{\omega i}$ and $V_{\omega i}$ are the matrixes of eigenvalues and eigenvectors of the matrix $COV_{\omega i}$. The diagonal term of the $D_{\omega i}$ consists of the variances $\sigma^2_{x\omega il}$, $\sigma^2_{y\omega il}$ and $\sigma^2_{z\omega il}$. The matrix $V_{\omega i}$ is used to transform the input vector from the initial coordinate system xyz to the local coordinate system $x_{\omega il}y_{\omega il}z_{\omega il}$.

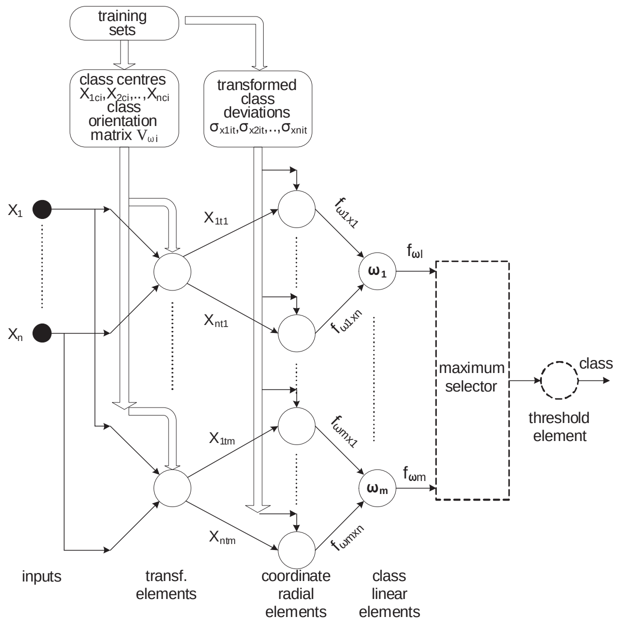

Figure 4. Classifier with decomposing RBEs (CDRBE2), which takes into account the orientation of the class axes of inertia

The architecture of the classifier [13], which corresponds to the approach described above, is presented in *figure 4*. There is a new layer including transforming elements, which realize the transformation of the input vector into the class local coordinate systems. The biases of the radial neurons in the second layer correspond to the standard deviations $\sigma_{x\omega il}$, $\sigma_{y\omega il}$ and $\sigma_{z\omega il}$.

The CDRBE2 can approximate the classes with shape, which can vary from n-dimensional ellipsoid (including n-dimensional sphere) to the shape, which is similar to the n-dimensional parallelepiped.

2.3. Classification of Multidimensional Patterns into Overlapping classes

Variant 4. (CRBEP). It is well-known, that the optimal solution of the task concerning classification into overlapping classes can be obtained on the basis of the Bayesian classification rule [7,10,15,20]. Its application requires the conditional probabilities $P(X/\omega_i)$, as well as the prior probabilities $P(\omega_i)$ to be known. The condition for correct classification of an input vector X into the class ω_i can be given by the inequality:

(7) $P(X/\omega_i)\,P(\omega_i) > P(X/\omega_j)\,P(\omega_j)$. $j=1\ldots m$, $j\neq I$,

where m is the number of classes.

This variant of the Bayesian classifier creates boundaries between classes, but it can not approximate the class areas. Furthermore, if the ratio of the number of elements in training sets is different from the ratio of the number of elements in real classes, the Bayesian classifier loses its optimality.

Let us suppose that the requirement for class areas approximation is valid and the ratio of numbers of elements in classes is changing during the classification of unknown sample. Under this formulation the following approach is realized [13]. The class areas are approximated using standard (or decomposing) RBEs and the accumulated during classification number of vectors of each of the classes is interpreted as class potential $V_{\omega i}$. This potential is defined as:

(8) $V_{\omega i} = N_i/N_{max}$

where N_i is the number of vectors, which are currently classified in the ω_i class; N_{max} is the number of vectors, which are currently classified in the class with maximum classifications.

As it is shown in *figure 9,* the class potential $V_{\omega i}$ introduces an additional correction $\Delta f_{\omega i} = k_v.V_{\omega i}/D(X,C_i)$ (where k_v is a weight coefficient and $D(X,C_i)$ is the Euclidean distance between the input vector X and the i-th class center) of the assessment $f_{\omega i}$, formed by the i-th RBE. This correction displaces the probability density function $f_{\omega i}$ with a value, which is proportional to the current number of vectors of the class ω_i.

The correction $\Delta f_{\omega i}$ can also be introduced as a correction of the RBEs biases. The i-th RBE bias can be given by the equation:

(9) $bi = 0{,}833/(k\sigma_{\omega i} + k_v.V_{\omega})$

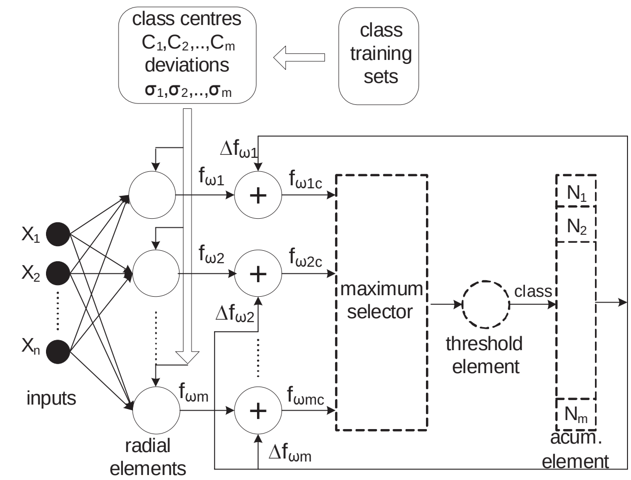

Figure 5. Classifier architecture (CRBEP), which takes into consideration the class potentials

The effect of the correction comes down to a displacement of the boundary between the overlapping class areas. The displacement depends on the ratio of the accumulated number of vectors in each of the classes.

2.4. Classifier Validation

As it was pointed out in section 2.1, it would be more efficient the class $\omega_1, \omega_2 \ldots \omega_m$ areas to be approximated by the classifier and all vectors, that do not get into these limited areas, to be associated with the class ω_s. That is why the classifiers presented above include one or set of RBEs corresponding to the classes $\omega_1, \omega_2 \ldots \omega_m$. For the ω_s class the approximating elements are missing. The goal of the validation is to obtain the optimal classifier parameters (k and k_v) for specific classification tasks.

In comparison with the standard cross-validation approach (K fold cross-validation), the validation approach proposed is based on the following procedure. Although the classifier creates models of the ω_1, $\omega_2 ... \omega_m$ classes, some elements of the ω_s class are used in classifier validation. This leads to the limitation of the class area dimensions, which is a precondition for a big part of elements of the ω_s class to be rejected from the classifier. In that case this result is a correct classification. For example, if we recognize color class of grain sample elements, we can include data about color characteristics of non grain impurities. Including this data, we limit the dimensions of the class areas. This leads to more correct recognition of grain sample elements**.**

3. Results and Discussion

The developed classifiers CSRBE, CDRBE1, CDRBE2 and CRBEP are simulated in MATLAB environment. Simulated data and experimental data about surface color and texture characteristics of the grain sample elements (extracted from object spectra) are used for classifiers validation, training and testing.

3.1. Classification through Class area Approximation Using CSRBE, CDRBE1, CDRBE2 and Simulated Data

The approximation of two classes ω_1 and ω_2 with normal distributed vectors, whose axes are parallel to the axes of the input vector coordinate system, is presented in *figure 6*. The number of the vectors in the two classes is respectively 1000 and 500. The approximation of two classes ω_1 and ω_2 with normal distributed vectors, whose axes have a random orientation in the feature space, is presented in *figure 7*.

Classification errors e_i and classification error rate e_o using CSRBE, CDRBE1 and CDRBE2 are shown in *table 1*. The

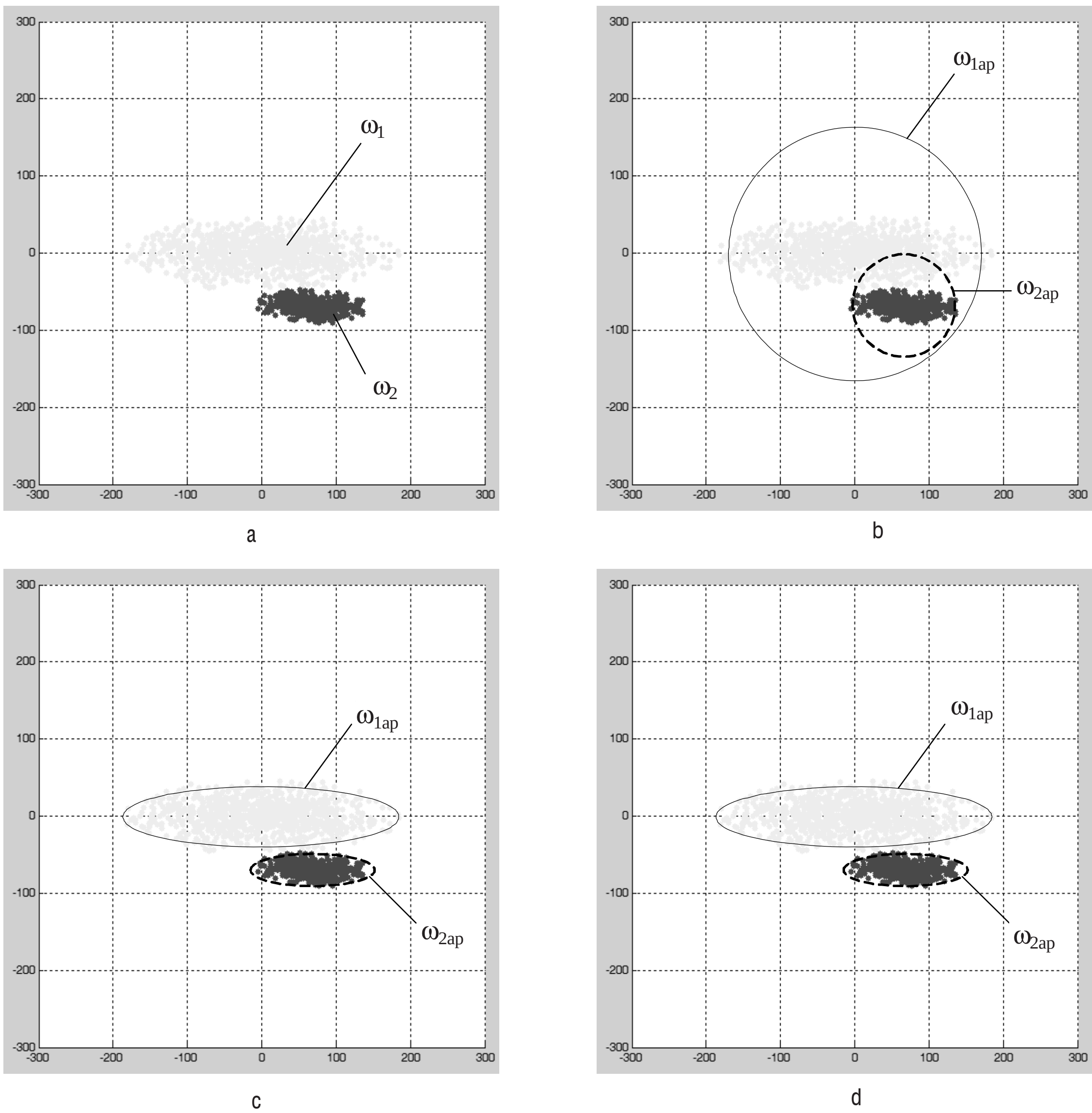

Figure 6. Approximation of classes ω_1 and ω_2 using CSRBE, CDRBE1 and CDRBE2. The class axes are parallel to the axes of the input vectors coordinate system: a)real class ω_1 and ω_2 areas; b) CSRBE approximation of the classes; c) CDRBE1 approximation of the classes; d) CDRBE2 approximation of the classes

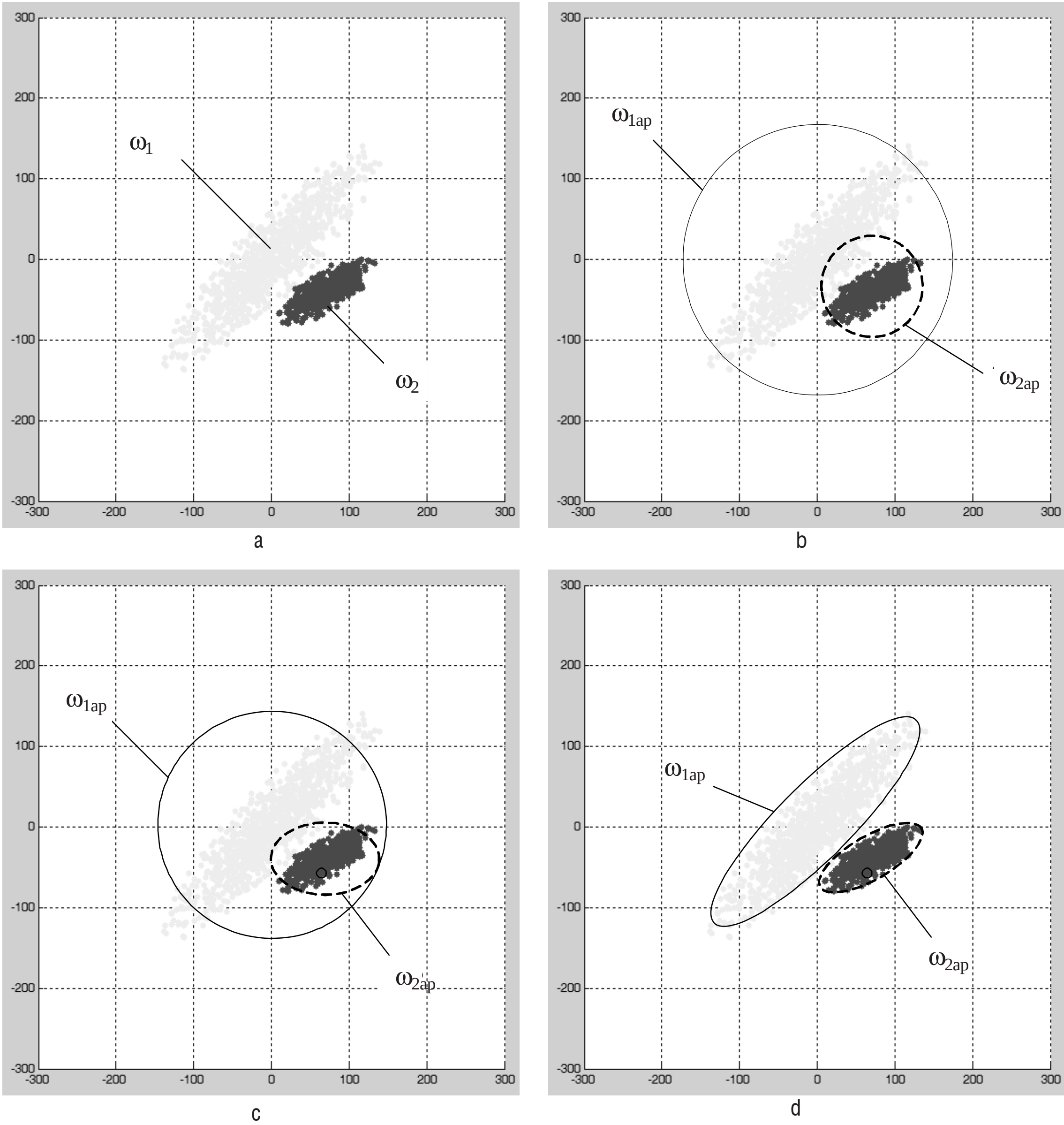

Figure 7. Approximation of classes ω_1 and ω_2 using CSRBE, CDRBE1 and CDRBE2. The classes have a random orientation in the input vectors coordinate system: a)real class ω_1 and ω_2 areas; b) CSRBE approximation of the classes; c) CDRBE1 approximation of the classes; d) CDRBE2 approximation of the classes

Table 1. Classification results obtained by CSRBE, CDRBE1 and CDRBE2

Class features	$\sigma^2_{\omega11}$=50, $\sigma^2_{\omega12}$=50, $\sigma^2_{\omega21}$=50, $\sigma^2_{\omega22}$=50, $C_{\omega1}$=(-100;0), $C_{\omega2}$=(100;0)			$\sigma^2_{\omega11}$=50,$\sigma^2_{\omega12}$=50, $\sigma^2_{\omega21}$=50,$\sigma^2_{\omega22}$=10, $C_{\omega1}$=(0;0), $C_{\omega2}$=(0;-120)			$\sigma^2_{\omega11}$=50,$\sigma^2_{\omega12}$=50, $\sigma^2_{\omega21}$=50,$\sigma^2_{\omega22}$=10, $C_{\omega1}$=(0;0), $C_{\omega2}$=(120;-120)			$\sigma^2_{\omega11}$=50,$\sigma^2_{\omega12}$=10, $\sigma^2_{\omega21}$=30,$\sigma^2_{\omega22}$=10, $C_{\omega1}$=(0;0), $C_{\omega2}$=(30;-40)		
Class orientation	$\Theta_{\omega11}$=0 deg $\Theta_{\omega21}$=0 deg			$\Theta_{\omega11}$=0 deg $\Theta_{\omega21}$=0 deg			$\Theta_{\omega11}$=45 deg $\Theta_{\omega21}$=45 deg			$\Theta_{\omega11}$=45 deg $\Theta_{\omega21}$=45 deg		
Errors,%	e_1	e_2	e_0	e_1	e_2	e_0	e_1	e_2	e_0	e_1	e_2	e_0
CSRBE-ref. classifier	1.59	0.64	1.27	9.86	0.0	6.57	1.27	4.67	3.00	4.98	19.83	9.93
CDRBE1	1.47	0.84	1.27	0.85	1.27	0.98	1.18	5.52	2.63	5.08	17.27	9.14
CDRBE2	1.27	0.68	1.07	0.64	1.27	0.85	0	0	0	0	0	0

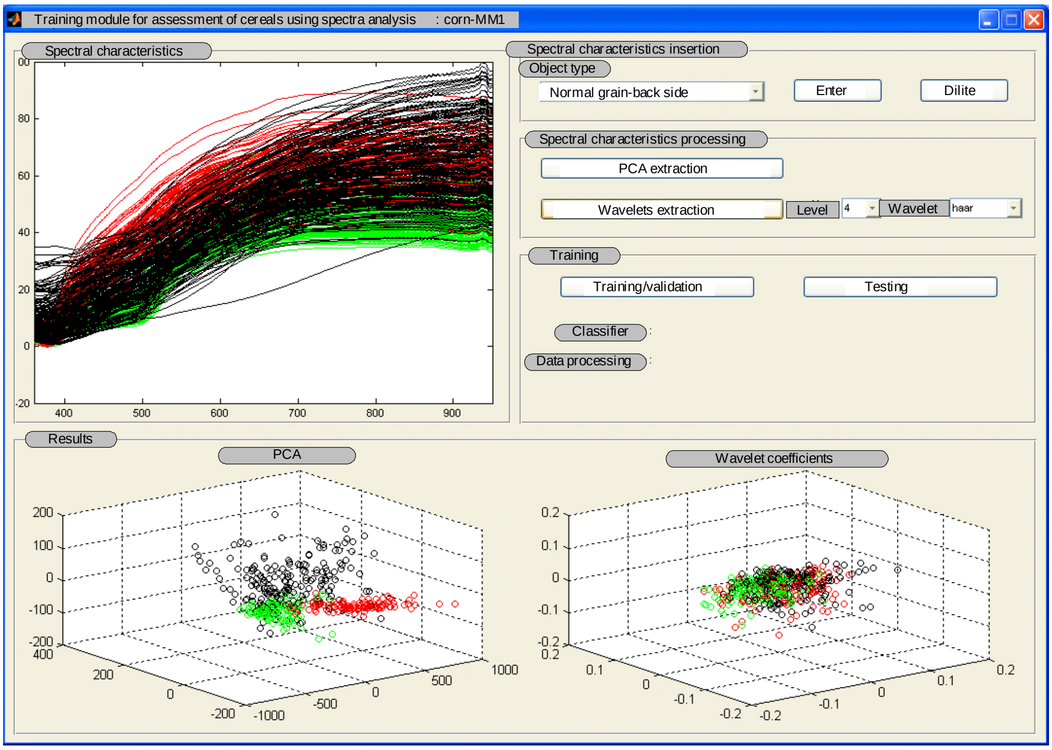

Figure 8. Presentation of the spectral characteristics and the results from their analysis

results presented are the average values obtained from 10 simulated testing samples with the same distribution parameters. The following denotations are used: $\sigma^2_{\omega 11}$, $\sigma^2_{\omega 12}$, $\sigma^2_{\omega 21}$ and $\sigma^2_{\omega 22}$ are the class variances in the directions of the inertial axes of the classes ω_1 and ω_2 respectively; $C_{\omega 1}$ and $C_{\omega 2}$ are the centers of the two classes; $\theta_{\omega 11}$ and $\theta_{\omega 12}$ are the angles between the main class axes and the x axis.

The classification errors are calculated using the equations

(10) $e_i = FP_i/(TP_i+FP_i)$.

e_i gives the relative part of objects from some class ω_i, which are assigned incorrectly to other classes $k=1...N$, where FP_i is the number of elements from the i-th class, which are incorrectly classified in other classes; TP_i is the number of correctly classified elements from the i-th class.

(11) $e_0 = \Sigma FP_i /(\Sigma TP_i + \Sigma FP_i)$.

e_0 (classification error rate) gives the relative part of all wrongly classified objects, were N is the number of classes.

Analysis of the results. The results obtained under discussed circumstances show, that the accuracy of class areas approximation influences sufficiently on the classification accuracy. Because of the CDRBE2 can better adapt to the shape and orientation of real class areas, the expectation that it will assure better classification accuracy is confirmed. For example, if we consider the ellipsoid shaped classes, whose main axes of inertia have orientation $\Theta_{\omega 11}$= 45 deg and $\Theta_{\omega 21}$ = 45 deg respectively, the CDRBE2 decreases the error e_0 with 9.93% and 9.14% in comparison with CSRBE and CDRBE1. This can be explained by the fact, that the CDRBE2 makes more precise approximation of real class areas. The errors of the three classifiers are approximately the same, when the classes are round shaped.

3.2. Classification of Maize Grain Sample Elements Using Features Extracted from Spectra

The spectral characteristics of grain sample elements are obtained using QE65000 spectrophotometer. Each characteristic is a vector with about 1500 components. The spectral characteristics are shown in the field „Spectral characteristics" (*figure 8*).

Principle Component Analysis (PCA) and combination of Wavelet descriptions and PCA are used for extracting typical features from object spectra and for reducing input data dimensionality. The following Wavelet coefficients are used: Wavelet1 - detailed coefficients and Wavelet2 - approximating coefficients. There is a possibly one of the following wavelet functions to be selected: Haar, Daubechies2, Coiflet2, and Symlet2. The level of decomposition can vary from $m=1$ to $m=4$. The most informative wavelet coefficients are chosen using PCA method. The well known PCA method is used as a reference method for feature extraction from spectra and for input data dimensionality reduction.

Three classes are presented in the field „Results" through the first 3 principle components or Wavelet coefficients.

The tree classifiers (CSRBE, CDRBE2 and CRBEP) used for object recognition on the basis of their spectral characteristics, are validated, trained and tested with the sets, presented in *table 2*.

	1cc	2cc	3cc	5cc	7cc	8cc	9cc
Training/ Validation	120	120	80	53	192	42	607 (val.)
Testing	30	30	20	13	48	11	152

Table 3. Classification results obtained, when elements from class 9cc are excluded from validating procedure and including in testing set

	Training errors, %									Testing errors, %		
Data model	PCA ref. data model			Wavelet1+PCA			Wavelet2+PCA			Wavelet2+PCA - selected		
Classifier	CDRBE2	CSRBE	CRBEP	CDRBE2	CSRBE	CRBEP	CDRBE2	CSRBE	CRBEP	CDRBE –selected		
	e_0	e_0	e_0	e_0	e_0	e_0	e_0	e_0	e_0	g_i	e_i	e_0
1cc										15.79	46.67	
2cc										0	50	
3cc										25	80	
5cc	0.82	47.78	0.82	0.33	47.78	0.33	0.82	9.88	0.82	0	61.54	27.62
7cc										3.33	39.58	
8cc										0	63.64	
9cc										36.49	0	

Table 4. Classification results obtained, when elements from class 9cc are included in validating and testing sets

	Training errors, %									Testing errors, %		
Data model	PCA ref. data model			Wavelet1+PCA			Wavelet2+PCA			Wavelet1+PCA - selected		
Classifier	CDRBE2	CSRBE	CRBEP	CDRBE2	CSRBE	CRBEP	CDRBE2	CSRBE	CRBEP	CDRBE –selected Val. Error=9.27%		
	e_0	e_0	e_0	e_0	e_0	e_0	e_0	e_0	e_0	g_i	e_i	e_0
1cc										0	16,67	
2cc										0	6,67	
3cc										10,53	15	
5cc	6.75	71.99	7.25	6.26	71.99	6.42	10.33	47.78	10.38	0	30,77	7.34
7cc										11,54	4,17	
8cc										0	18,18	
9cc										9,03	2,24	

The classification results of grain sample elements using the three classifiers and the three data models are presented in *table 3* and *table 4*. The results in the first table are obtained, when elements from class 9cc are excluded from the validation procedure. The results in *table 4* are obtained, when some elements from class 9cc are included in the validation.

The classification error g_i is calculated using the equation:

(12) $g_i = FN_i/(TP_i + FN_i)$.

g_i gives the relative part of objects from other classes, which are assigned to class I, where FN_i is the number of elements from other classes, which are assigned to the i-th class.

Analysis of the results. The results obtained confirm the effectiveness of the proposed classification approach, classifier architectures, data models and validation approach. For example, the testing error e_0 of the selected classifier (CDRBE2) and data model (Wavelet+PCA) is 27.62%, if elements from class 9cc are excluded from the validation procedure. This error is 7.34%, when some elements from class 9cc are included in the validation. The validation approach proposed decreases the testing error with 20.3% under specific classification circumstances.

The choice of the appropriate classifier for specific classification task has significant influence on the classification results. For example, the training errors obtained using CDRBE2, CSRBE and CRBEP classifiers and PCA data model, are 6.75%, 71.99% and 7.25% respectively.

3.3. Classification into Overlapping Classes Using CSRBE, CRBEP2 and Bayesian Classifier

The CSRBE, CRBEP and Bayesian (reference classifier) classifiers are used for vector classification into overlapping

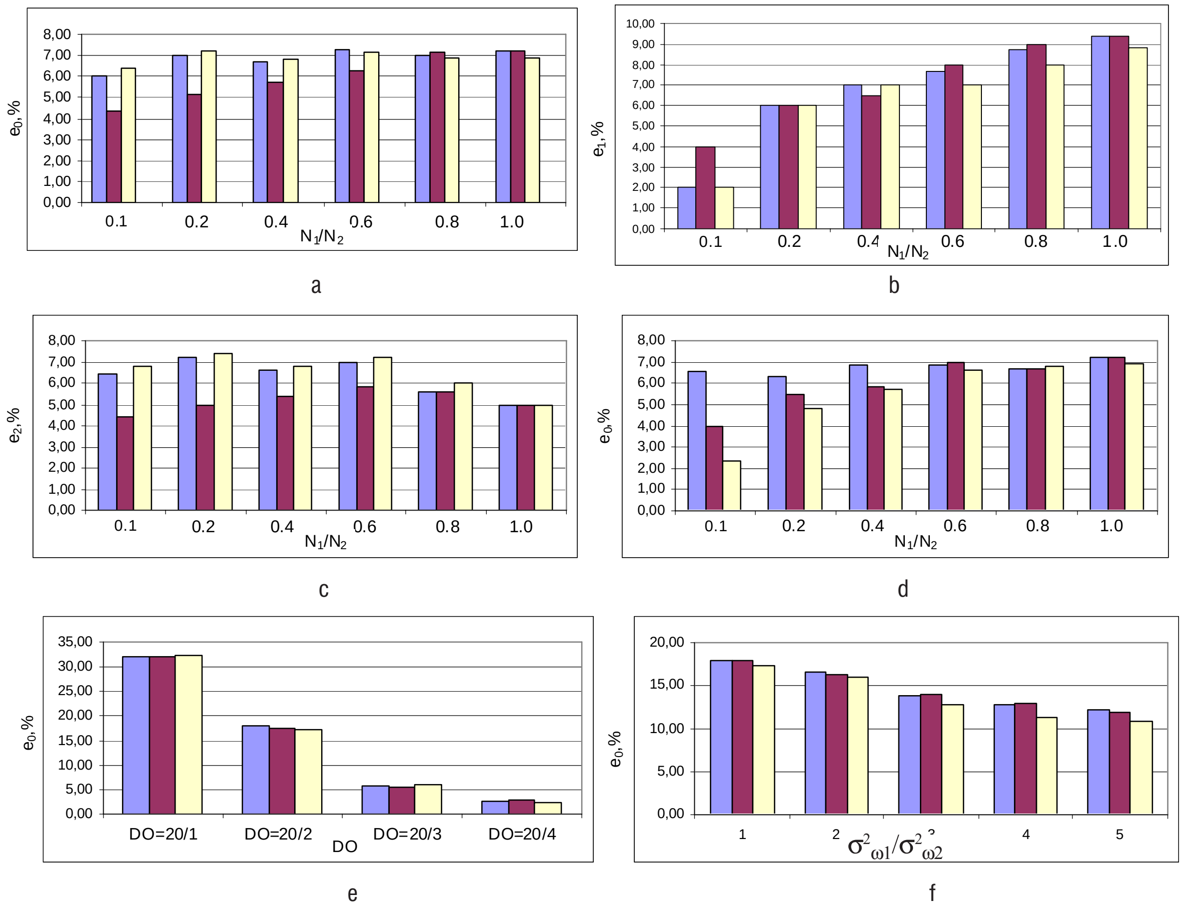

Figure 12. Comparison between the classification accuracy, obtained by the CSRBE, CRBEP2 and Bayesian classifier: a) $e_0 = f(N_1/N_2)$, $P(\omega_1)/P(\omega_2)=1\neq N_1/N_2$; b) $e_1 = f(N_1/N_2)$, $P(\omega_1)/P(\omega_2)=1\neq N_1/N_2$; c) $e_2 = f(N_1/N_2)$, $P(\omega_1)/P(\omega_2)=1\neq N_1/N_2$; d) $e_0 = f(N_1/N_2)$, $P(\omega_1)/P(\omega_2)=N_1/N_2$; e) $e_0 = f(DO)$, $P(\omega_1)/P(\omega_2)=N_1/N_2$; f) $e_0 = f(\sigma^2_{\omega 1}/\ \sigma^2_{\omega 2})$, $P(\omega_1)/P(\omega_2)=N_1/N_2$

-CSRBE, -CRBEP, -Bayes classifier

classes. Two classes ω_1 and ω_2 with two-dimensional vectors are considered. The classifiers are validated and trained using sets, which include N_{1t} and N_{2t} vectors respectively. They are tests with sets, which include N_1 and N_2 vectors. The classification results are presented in *figure 12*, where: $P(\omega_1)$ and $P(\omega_2)$ are the priori probabilities of the two classes (they correspond to N_{1t} and N_{2t}); $\sigma^2_{\omega 1}$ and $\sigma^2_{\omega 2}$ are the class variances; DO represents the class overlap degree, which is determined by the following equality:

(13) $$DO = (\sigma^2_{\omega 1} + \sigma^2_{\omega 2})/\ D_{12},$$

where D_{12} is the Euclidean distance between the class centers.

Analysis of the results. In comparison with the Bayesian classifier, the CRBEP classifier assures better classification accuracy, when the ratio of the number of elements in testing set (N_1/N_2) is different from the ratio of the number of elements in training set (N_{1t}/N_{2t}). N_{1t} and N_{2t} define the priory probabilities $P(\omega_1)$ and $P(\omega_2)$. This result confirms the expectation, that the Bayesian classifier loses its optimality under such circumstances. Since the CRBEP takes into account the number of elements in each of the classes after each classification, it assures better accuracy than the Bayesian reference classifier (*figure 12a*). The difference between the errors of the two classifiers increases, when the difference between the two ratios increases. For example, if $N_1/N_2 = 0.1$, the CRBEP decreases the classification error e_0 with 2% by comparison with the Bayesian classifier. If the values of the two ratios are nearly equal, then the errors of the two classifiers are similar. It is relevant to remark that CRBEP increases the accuracy of the class, which has bigger number of elements (*figure 12c*). This is not true for the class, whish has a smaller number of elements (*figure 12b*).

The Bayesian classifier assures the biggest accuracy (*figure 12d*), when the ratios N_{1t}/N_{2t} and N_1/N_2 are equal, especially when the values of these ratios are small. For example, if $N_1/N_2 = 0.1$, the Bayesian classifier's error is with 1.64% and 4.18% smaller than the errors obtained by the CRBEP and CSRBE respectively. The difference between the errors of the three classifiers decreases, when the ratio N_1/N_2 decreases.

The classification accuracy of the three classifiers decrease, when the degree of overlap of class areas increases (*figure 12e*). For example, if DO=200/40, the errors of the three classifiers are about 2.5%. This error is 17.6%, when DO=200/10.

The ratio of the class variances $\sigma^2_{\omega 1}/\ \sigma^2_{\omega 2}$ influences sufficiently on the classification accuracy. The accuracy of the three classifiers increases, when this ratio increases (*figure 12f*).

4. Conclusions

The results from the investigation can be summarized as follows:

1. The task for multidimensional patterns classification in classes, which are located in a noisy environment, can be effectively solved through class areas approximation using classifiers, based on RBEs.

2. The results obtained under discussed circumstances show, that the accuracy of class areas approximation influences sufficiently on the classification accuracy. The CDRBE2 can better adapt to the shape and the orientation of real class areas and it assures better classification accuracy in comparison with the CSRBE and CDRBE1 classifiers.

3. These results are confirmed, when we recognize the color class of grain sample elements. The selection of a proper classifier has a significant influence on the training and testing classification accuracy. The CDRBE2 classifier decreases training and testing errors with 6.22% and 66.92% in comparison with the CSRBE reference classifier.

4. In comparison with the Bayesian classifier, the CRBEP assures better classification accuracy, when the ratio of the number of elements in testing set is different from the ratio of the number of elements in training set. This result confirms the expectation, that the Bayesian classifier loses its optimality under such circumstances. Since the CRBEP takes into account the number of elements in each of the classes after each classification, it assures better accuracy, than the Bayesian classifier. The difference between the errors of the two classifiers increases, when the difference between the two ratios increases. If the values of the two ratios are nearly equal, then the errors of the two classifiers are similar.

5. The classification accuracy of the three classifiers decreases, when the degree of overlap of the class areas increases.

6. The investigation carried out shows, that if the validation procedure is based on the training sets of classes with compact areas only, the classes areas are increased and a big part of noisy vectors are associated with some of these classes. Better results are got, if noisy vectors are included in classifier validation. For example, the classification error e_0 increases with 20.3%, when data about non grain impurities is excluded from validation procedure.

Acknowledgements

This investigation is part of research activities in the context of the research project „Development of intelligent technologies for assessment of quality and safety of food agricultural products" funded by the Bulgarian National Science Fund.

References

1. Brosnan, T., D. W. Sun. Inspection and Grading of Agricultural and Food Products by Computer Vision Systems - a Review. - *Computers and Electronics in Agriculture*, 36, 2003, 193-213.

2. Buchtala, O., M. Klimek, B. Sick. Evolutionary Optimization of Radial Basis Function Classifiers for Data Mining Applications. - *IEEE Transactions on Systems, Man, and Cybernetics,* Part B: Cybernetics, 35, 2005, Iss. 5, 928-947.

3. Cristianini, N., J. Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press, 2006.

4. Gonzalez, J., I. Rojas, J. Ortega, H. Pomares, F. J. Fernandez, A. F. Diaz. Multiobjective Evolutionary Optimization of the Size, Shape, and Position Parameters of Radial Basis Function Networks for Function Approximation. - I*EEE Transactions on Neural Networks*, 14, 2003, Iss. 6, 1478-1495.

5. Hammond, M. H., C. J. Riedel, S. L. Rose-Pehrsson, F. W. Williams. Training Set Optimization Methods for a Probabilistic Neural Network. - *Chemometrics and Intelligent Laboratory Systems*, 71, 2004, Is. 1, 73-78.

6. Hsu, C. W., C. J. Lin. A Comparison of Methods for Multiclass Support Vector Machines. - I*EEE Transactions on Neural Networks*, 13, 2002, Iss. 2, 415-425.

7. Huang, H. J., C. N. Hsu. Bayesian Classification for Data From the Same Unknown Class. - *IEEE Transactions on Systems, Man, and Cybernetics*, Part B: Cybernetics, 32, 2002, Iss. 2, 137-145.

8. Kotsiantis, S., P. Pintelas. Increasing the Classification Accuracy of Simple Bayesian Classifier. Lecture Notes in Artificial Intelligence, AIMSA 2004, Springer-Verlag, 3192, 2004, 198-207.

9. Kotsiantis, S., P. Pintelas. Recent Advances in Clustering: A Brief Survey, WSEAS Transactions on Information Science and Applications, 1 (1), 2004, 7, 3-81.

10. Kuncheva, L. I. On the Optimality of Naïve Bayes with Dependent Binary Features. - *Pattern Recognition Letters*, 27, 2006, Iss. 7, 830-837.

11. Lee, Y., Y. Lin, G. Wahba. Multicategory Support Vector Machines. - *Journal of the American Statistical Association*, 99 (465), 2004, 67-81.

12. Mladenov, M., M. Dejanov. Application of Neural Networks for Seed Germination Assessment. Proceedings of the 9th WSEAS International Conference on Neural Networks CNN'08, Sofia, Bulgaria, 2008, 67-72.

13. Mladenov, M., S. Penchev, M. Dejanov, M. Mustafa. Quality Assessment of Grain Samples Using Color Image Analysis. Proc. f he 8th IASTED International Conference on Signal Processing, Pattern Recognition and Applications (SPPRA 2011), Innsbruck, Austria, 2011, 37-44.

14. Nam, M. D., T. C. Thanh. Approximation of Function and its Derivatives Using Radial Basis Function Networks. - *Applied Mathematical Modelling*, 27, 2003, Iss. 3, 197-220.

15. Neville, J., D. Jensen, B. Gallagher. Simple Estimators for Relational Bayesian Classifiers. Proc. of the Third IEEE International Conference on Data Mining ICDM'2003, 2003, 609-612.

16. Nearest-Neighbor Methods in Learning and Vision. Edited by Shakhnarovish, Darrell and Indyk. The MIT Press, 2005.

17. Oyang, Y. J., S. C. Hwang, Y. Y. Ou, C. Y. Chen, Z. W. Chen. Data Classification with Radial Basis Function Networks Based on a Novel Kernel Density Estimation Algorithm. - I*EEE Transactions on Neural Networks*,16, 2005, Iss. 1, 225-236.

18. Patra, P. K., M. Nayak, S. K. Nayak, N. K. Gobbak. Probabilistic Neural Network for Pattern Classification. Proc. of the International Joint Conference on Neural Networks IJCNN '02, 2002, 1200-1205.

19. Peng, S. Y., P.E. Hasler, D. V. Anderson. An Analog Programmable Multidimensional Radial Basis Function Based Classifier. - I*EEE Transactions on Circuits and Systems*, 54, 2007, Iss. 10, 218-221.

20. Ritter, G., M. T. Gallegos. A Bayesian Approach to Object Identification in Pattern Recognition. Proc. of the 15th International Conference on Pattern Recognition, Barcelona, Spain, 2, 2000, 418-421.

21. Rutkowski, L. Adaptive Probabilistic Neural Networks for Pattern Classification in Time-varying Environment. - I*EEE Transactions on*

Neural Networks, 15, 2004, Iss. 4, 811-827.

22. Sarimveis, H., P. Doganis, A. Alexandridis. A Classification Technique Based on Radial Basis Function Neural Networks. - *Advances in Engineering Software*, 37, 2006, Iss. 4, 218-221.

Manuscript received on 21.02.2011

Mirolyub Mladenov *received MSc degree in Automatics from the Technical University of Sofia, Bulgaria, in 1974. He obtained his PhD degree in Automation in 1985. He is currently an Associated Professor at the University of Ruse „A. Kanchev“, Bulgaria. He is head of the Department of „Automatics, Information and Control Engineering“. His research interests are in the Smart Sensors, Computer Vision, Image Processing, Pattern Recognition and Neural Networks. He has published about 80 papers, 4 books and a monography in this area.*

Contacts:
Department of Automatics, Information and Control Engineering
University of Rouse, 8 Studentska Str., 7017 Ruse, Bulgaria
tel. +359 82 888 747, fax: +359 82 845 708
e-mail: mladenov@uni-ruse.bg, http://www.uni-ruse.bg