

Grain Sample Quality Assessment using Color Image and Spectra Analysis

Miroljub Mladenov, Martin Deyanov
University of Ruse, Bulgaria

Roumiana Tsenkova
University of Kobe, Japan

Metin Mustafa, Stanislav Penchev, Tsvetelina Draganova
University of Ruse, Bulgaria

ABSTRACT: The paper presents the approaches, methods and tools for assessment of main quality features of grain samples using analysis of objects color images and spectral characteristics, which are implemented in the frames of INTECHN project. Visible features like color, shape, and dimensions are extracted from the object images. Information about object color and surface texture is obtained from the object spectral characteristics. The categorization of grain sample elements in three quality groups is accomplished using two data fusion approaches. The first approach is based on the fusion of the results about object color and shape characteristics obtained using image analysis only. The second approach fuses the shape data obtained using image analysis and the color and surface texture data obtained by spectra analysis. The results obtained by the two data fusion approaches are compared.

KEYWORDS: grain sample quality assessment, color image analysis, spectra analysis, classification, data fusion.

Introduction

The assessment of the food quality and safety is an important part of the food production chain. The higher food quality requirements demand the development of new, objective technologies for food quality assessment. In conformity with these requirements is the main goal of INTECHN project “Development of

Intelligent Technologies for Assessment of Quality and Safety of Food Agricultural Products” founded by the Bulgarian National Science Fund. Computer based, express, objective, intelligent technologies, methods and tools for assessment of the main quality and safety features of different food products like cereals, fruits, vegetables, milk, dairy products, meat, meat products and eggs were developed and investigated within the frames of the project.

Two different approaches are used to obtain complex assessment of food products quality features. The first approach is based on analysis of object color images using a computer vision system. Visible features related to the color characteristics, shape and dimensions could be obtained because of such analysis. The second approach that uses analysis of object spectral characteristics can form information about object features related to the object surface color and texture characteristics. Fusing the results from two kinds of analysis, we can get more complete and accurate notion about main quality features.

This paper discusses the INTECHN platform approaches, methods and tools for analysis of color images and spectral characteristics of the grain sample elements, as well as the approaches for the fusion of the results from the two kinds of analysis. These approaches are illustrated in the assessment of main quality features of maize grain samples. Similar methods are used for the assessment of the quality of the other food products.

According to the Bulgarian National Standards, the main quality features of maize grains are inherent for the variety appearance, shape, color, smell, taste, moisture, and impurities. Whole grains and broken grains (bigger than the half of the whole grain) which have inherent for the variety appearance, shape, and color are considered as normal grains. The grain impurities include broken grains (smaller than the half of the whole grain), heat-damaged grains, burned grains, small grains, shriveled grains, green grains, sprouted grains, and mouldy grains. The non-grain impurities include corncob particles, leaf and stem fractions, pebbles, soil, sand, and metal particles. Smutty grains, infected (with *Fusarium*) grains, as well as harmful elements (bunt) are included in the last grain sample group.

The grain quality characteristics specified above (excluding smell, taste, and moisture) are evaluated by an expert based on its visual estimation. They are mainly related to the appearance, color characteristics, shape, and dimensions of the sample elements. The main trend in the last few years is to use Computer Vision Systems (CVS) in order to evaluate such characteristics [BS03, M+11]. To obtain a complex assessment of the grain quality using data about color characteristics, shape and dimensions of the grain sample elements, is a complicated and multilevel task. This is

because the color characteristics shape and dimensions of the elements in a sample vary within a wide range.

Many results are published, in which different quality features, like authenticity [L+93, JPV00, P+03], variety [LJS99, MJ00], infections [N+88, M+11], purity [M+02, M+04], germination [MD08], mechanical damages [M+11] and other quality features are identified on the basis of color image analysis.

Some preliminary investigations [M+11] show that the analysis of object color images could not give sufficiently precise assessment of some grain sample elements like infected grains, mouldy grains and non grain impurities. This is determined by the fact that in comparison with the normal grains, not only color features of such objects are changed, but the surface texture is changed too. It is difficult to detect this change using a CVS.

That is why we expect a more accurate assessment of these sample elements to be obtained using spectra analysis. Unfortunately, information about object shape and dimensions cannot be extracted from spectra. To categorize the grain sample elements into normative groups specified above a fusion is performed for shape data obtained by image analysis and color and texture data obtained using spectra analysis.

Visible (VIS) and near infrared (NIR) spectra analyses are applied in the assessment of different grain quality features [H+08]. They are mainly used in tasks, related to the determination of qualitative and quantitative features like grain composition [BPS06], dry matter content [C+06], moisture content [M+10], starch [S+05], protein [W+04], glutenin [Car03], vitamins and toxins [F+09], vitamin and mineral content [O+07], as well as for the detection of grain infections [P+10].

Different methods like Principal Component Regression, Partial Least Squares Regression, Principal Component Analysis (PCA), Hierarchical Cluster Analysis, and other methods are used for development of models to predict a property of interest, feature extraction, and large and complex data sets reduction.

Different approaches are applied for categorization of the investigated objects. Effective results are got using classifiers based on Artificial Neural Networks (ANN), discriminant analysis, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Soft Independent Modeling of Class Analogy (SIMCA). Some of these methods are used in commercial software products (Pirouette, Unscrambler) for different products quality assessment based on spectra analysis.

The main goal of this paper is to present the INTECHN platform approaches, methods and tools for assessment of main quality features of the grain sample elements using image and spectra analysis, as well as the results from data fusion of the two kinds of analysis.

1. Materials and Methods

1.1. INTECHN platform hardware

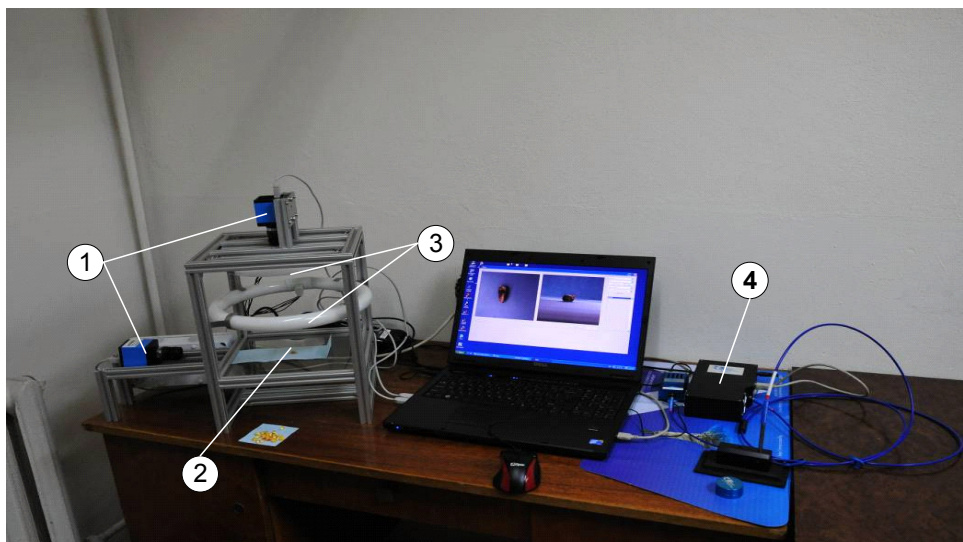


Fig. 1. INTECHN platform hardware components

The INTECHN hardware system (Fig.1) consists of the following main components: CVS and spectrophotometer (4) (QE65000, Ocean Optics, USA). The CVS includes two color CCD cameras (1) which give a possibility to form color images of investigated object (2) in two planes (horizontal and vertical). The illuminate system (3) is used for direct object illumination.

1.2. Grain sample classification groups

The characteristics of grain sample elements, which are on principle being evaluated by an expert, based on its visual estimation, are assessed using CVS within the frames of the INTECHN project. Groups (classes) and subgroups (subclasses) in which the elements of the maize grain sample have to be distributed are presented in Table 1.

Table 1. Maize grain sample classes and subclasses

Normative classes	Color classes	Shape classes
1cst - standard kernel (whole grains and broken grains bigger than the half of the whole grain,) with appearance, shape and color inherent for the variety	1cc- grains with color inherent for the variety, back side 2cc- grains with color inherent for the variety, germ side 3cc- heat-damaged grains	1csh- whole grains with inherent for the variety shape 2csh- broken grains bigger than the half of the whole grain
2cst-grain impurities: broken grains smaller than the half of the whole grain, heat-damaged grains, small grains, shriveled grains, green grains, sprouted grains, infected (with <i>Fusarium</i>) grains, smutty grains.	4cc- green grains 5cc- moldy grains 6cc- smutty grains	3csh- broken grains smaller than the half of the whole grain and small and shriveled grains
3cst- non grain impurities: corn-cob particles, leaf and stem fractions, pebbles, soil and sand, as well as harmful elements	7cc- infected (with <i>Fusarium</i>) grains 8cc- sprouted grains 9cc- non – grain impurities	4csh- non – grain impurities

It is evident from the table that the main determinative characteristics of grain sample elements are related to the object color and shape. Because of the fact that color and shape features are described in a different manner, it is advisably the assessment these characteristics to be made separately. After that, the results from two assessments have to be fused to obtain the result about object categorization to one of the normative classes. In terms of the classification procedure, color and shape groups are divided in several subgroups. Color object characteristics are separated in 8 basic classes corresponding to the typical for the different sample elements color zones and 1 additional class that corresponds to the non – grain impurities (it is impossible to define a compact class for non – grain impurities). According to the object shape, the objects are divided in 3 basic classes corresponding to the whole grains, broken grains bigger than the half of the whole grain and broken grains smaller than the half of the whole grain, and 1 additional class that correspond to the non – grain impurities. Each of the three basic shape classes is further divided in 6 shape subclasses.

1.3. Feature extraction from images

RGB, HSV, XYZ, and YCbCr color models are used for separating object area from the background and for extraction of different color zones inside the object area (inherent for the variety color, heat-damaged grains, green grains, smutty grains, infected (with *Fusarium*) grains, bunt, non – grain impurities). Furthermore, four color texture models [MD08, M+11] are developed for this purpose. It is expected they will better underline zone color ratios in the input RGB image.

To represent the shape of the investigated objects, 10 – dimensional vector descriptions are used [M+02, M+11]. It is typical for the maize kernels that its contour has a huge asymmetry at the part where the germ exists and at the opposite part. It is easy to locate the germ in the whole grains and to build contour descriptions and models with proper orientation. For the broken grains, depending on what part of the whole is remaining (with the germ or without it) the contour descriptions could be sufficiently different. That is why it is necessary for classes 2csh and 3csh (to be more precise – for their corresponding subclasses) to define two types of descriptions and models: for shapes where the germ exists in the remaining part of the grain and where it is not. All the shapes of the investigated objects are divided into 18 groups (subclasses).

1.4. Features extraction from spectra and data dimensionality reduction

The classification of grain sample elements based on their spectral characteristics is made in groups corresponding to the color classes presented in Table 1.

The spectral characteristics of grain sample elements are vectors with about 1500 components. PCA and combination of Wavelet descriptions and PCA are used for extraction of typical features from object spectra and for reduction of input data dimensionality. The following Wavelet coefficients are used: Wavelet1- detail coefficients and Wavelet2- approximation coefficients. The operator can select one of the following Wavelet functions: Haar, Daubechies2, Coiflet2, Symlet2. The decomposition can vary from 1 to 4. The most informative Wavelet coefficients are chosen using the PCA method.

1.5. INTECHN platform classification procedures

INTECHN platform uses specific classification strategy, classifiers, and validation approach. They are conditioned by the specificity of the classification tasks concerning grain sample element recognition.

Classification approach. If the classes (built from color, shape, PCA and Wavelet+PCA descriptions) are presented in the feature space, a part of them (1cc, 2cc,...8cc) will form comparatively compact class regions. The sets of descriptions extracted from the grain sample training sets are used for developing the models of these grain sample groups. Each class model is presented by the class centre (the average value of the class training data) and the class boundary surface. The boundary surface is determined through a threshold value of the covariance of the class training data. A correct model for the 9cc class could not be created because of the fact that the spectral characteristics of elements of this class could be sufficiently different in each subsequent grain sample.

As a correct model for the 9-th grain group could not be created, a part of the descriptions of such objects from the testing set could get into the boundaries of the other eight classes defined. A big part of them would get outside the class regions and could be located in a random place in the feature space. These descriptions could be considered as noisy vectors. It could be assumed that the comparatively compact class regions of the objects from the first eight groups are submerged in a noisy environment. Therefore, the task for categorization of the grain sample elements can be interpreted as a task for classification in classes, whose boundaries have definite shapes, dimensions and location in the feature space, and they are situated in a noisy environment.

Under this formulation, the use of popular strategies like LDA, CA, SVM, KNN and some other methods, which build boundaries between class regions, is obviously not a good choice. This is due to the fact that for the class 9cc, which correspond with the 9-th grain group, a correct model cannot be created.

The task for grain class modeling is reduced to a task for approximation of the grain class regions. For this purpose, classifiers based on Radial Basis Elements (RBEs) are used. Such classifiers are chosen in terms of the simplicity of the classification procedure and the accuracy of the class region 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 input vectors will be included within the class boundary and it is easy to change the dimensions of the particular class region.

INTECHN platform classifiers. The following classifiers [M+11] were used for class region approximation:

Classifier with standard RBEs (CSRBE). Only one RBE is used for approximation of each class region. The RBEs centers correspond to RGB average values of the pixels from training sets. The RBEs biases are set in correspondence with their standard deviations. The CSRBE approximates

round shaped classes only. The CSRBE can be considered as a referent classifier, because its architecture includes standard RBEs and it realizes the classification strategy described above.

Classifier with decomposing RBEs (CDRBE). The CDRBE classifier architecture includes three layers. The first layer consists of m (m is the number of classes) transforming elements, which recalculate input vectors coordinates in local coordinate systems whose axes coincide with class axes of inertia. The second layer consists of $n \times m$ RBEs that are distributed in m sub layers. The number of RBEs in each sub layer is equal to n (n is the input vector dimensionality). The RBEs centers coincide with average values of the projections of the training sets vectors onto corresponding coordinates of the class local coordinate system. The RBEs biases are set in correspondence with the standard deviations of these projections. The third layer consists of m RBEs. The weights of all RBEs are equal (1, 1... 1). The RBEs outputs are the weighted distances of input vector to the centers of non – spherical classes.

The CDRBE classifier gives a possibility to form classes whose dimensions along the directions of separate coordinate axes are different. Changing the RBEs biases and the threshold Δ we can vary the class shape from a sphere to the shape close to a parallelepiped.

Classifier with RBEs which takes into consideration the class potentials (CRBEP). The CRBEP classifier [M+11] approximates the class regions using standard (or decomposing) RBEs and takes into consideration the class potentials. The number of vectors in each of the classes, accumulated during the classification process, is interpreted as a class potential. It introduces an additional correction of the assessment formed by i -th RBE. The effect of the correction comes down to the displacement of the boundary between the two overlapping parts of class regions. The displacement depends on the ratio of accumulated number of vectors in each of the classes.

Validation approach. The classifiers presented in this section are trained and validated using a specific cross-validation procedure. The goal of the validation is to select the appropriate data model and classifier, as well as to obtain the optimal classifier parameters.

In comparison with the standard cross-validation approach (K fold cross-validation), the INTECHN platform validation is based on the following procedure. Although the classifiers create models of the first eight grain sample groups, some elements of the 9cc group are used in classifier validation. This leads to the limitation of the class area dimensions, which is a precondition for a big part of non-grain impurities to be rejected from the classifier. In that case this result is a correct classification.

1.6. Fusing of the results from the color and shape analysis

Because of the fact that color and shape features are described in a different manner, the assessment of these characteristics is made separately. After that, the results from the two assessments have to be fused in order to obtain the object's final categorization to one of the normative classes.

Different variants for fusing the results from color and shape analysis of grain sample elements are developed at different stages of the investigation. The algorithms developed could be associated with hierarchical clustering algorithms. Their typical feature is that different criteria for class merging are used at different levels of data fusion.

Variant 1. A comparatively simplified fusion scheme is used in the first algorithm. It is presented in Fig. 2. The input data (input classes) are separated in two groups – color characteristics data and objects shape data. The first group consists of 10 color zones inherent for the maize grains. The second group consists of 18 shape subclasses (1scsh, 2scsh... 18scsh). The first six of them correspond to different shape models of whole kernels, the next six – to models of broken grains bigger than the half of whole grain and the last six – to models of broken grains smaller than the half of the whole grain.

The color class (1cc, 2cc... 8cc) is determined based on preliminary defined combinations of color zones at the first stage of fusing the results from the color analysis. The shape subclasses are merged into one of the three main shape classes (1csh, 2csh, and 3csh).

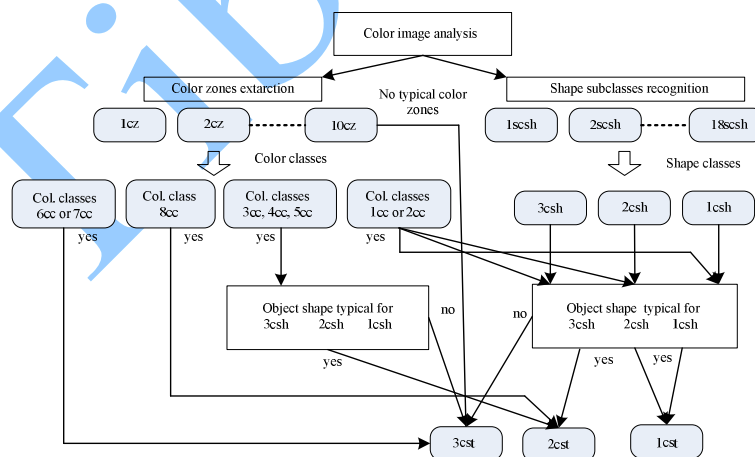


Fig. 2. Fusing the results from color characteristics and shape analysis. Variant 1

At the second stage of the analysis, the fusion of color and shape classes is made in order to form the final decision of object classification in

one of the three normative classes (1cst, 2cst, and 3cst). The assessment whether the shape of the object is typical for one of the three main grain classes or not is used as a fusion criterion for color classes 1cc to 5cc. For 6cc, 7cc and 8cc classes the shape is not important at all.

Variante 2. The second algorithm (Fig. 3) uses color and combined topological models of typical color zones. The topological models represent the plane distribution of the color zones within the object area. A set of color topology models (when 3 or more typical for the kernels color zones are found) and combined topology models (when only 2 typical color zones are found) is preliminary defined. The combined topological models represent the plane distribution of some shape element (kernel prickle or rear area) and the color zones found.

The final objects categorization when such topology is found is performed without taking into account the object's shape but only its area. The shape and area are the most important for the final categorization of the objects when only one typical color zone is found. When the object belongs to 6cc, 7cc or 8cc its shape is not important at all for taking the decision.

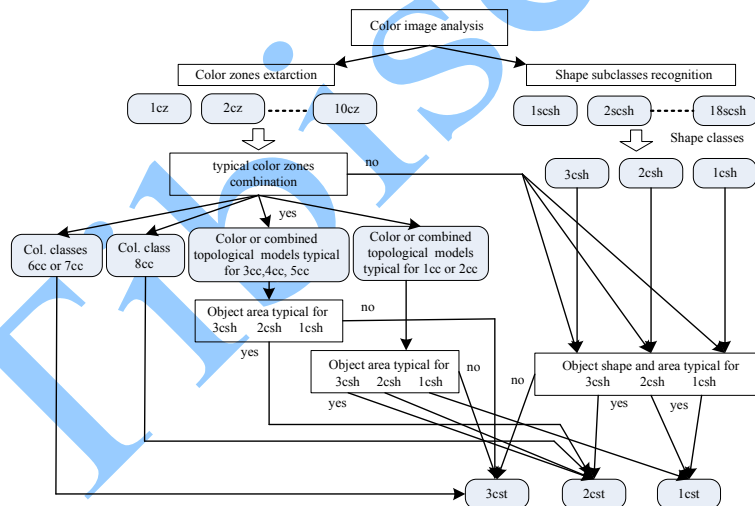


Fig. 3. Fusing the results from color characteristics and shape analysis. Variante 2

Variante 3. The third data fusion algorithm (Fig. 4) is based on data about object color characteristics obtained through analysis of their spectral characteristics and data about object shapes extracted from object images. This variant of color and shape data fusion is conditioned by the fact that the recognition of color class of grain sample elements using spectral data is more precise.

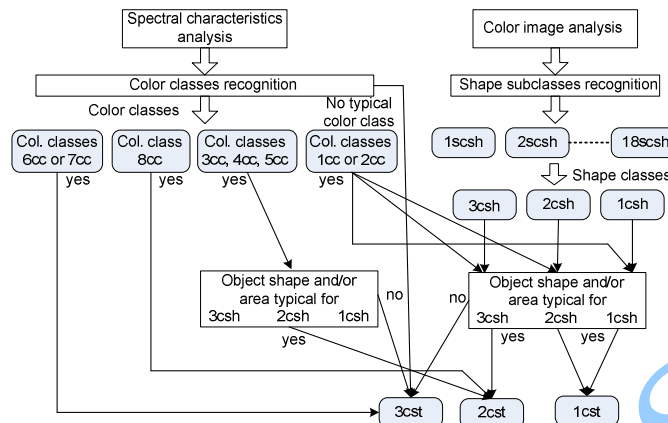


Fig. 4. Fusing the results from color characteristics and shape analysis. Variant 3

The main criterion for the final categorization is the object color class. The correspondence of object shape and/or area to the typical for the different shape classes shape and area is an additional criterion.

2. Results and discussion

2.1. Training and testing sets

The developed procedures for assessment of grain sample quality are validated, trained, and tested with sets presented in Table 2.

Table 2. Training and testing sets

Color class recognition							
Classes	1cc	2cc	3cc	5cc	7cc	8cc	9cc
Training sets	10	10	12	15	18	19	
Testing sets	47	81	44	24	74	39	168
Object shape classification							
Classes	1csh	2csh	3csh	4csh			
Training sets	120	135	135				
Testing sets	122	63	11	281			
Classification in normative classes							
Classes	1cst	2cst	3cst				
Testing sets	117	118	242				

2.2. Color characteristics analysis using CVS

Classification error rates of color class recognition using the classifier selected by the validation procedure are presented in Table 3. The “Testing errors1” represents the results when the group of non-grain impurities is excluded from the testing set, while the “Testing errors2” represent the results when this group is included in the testing set.

Table 3. Classification errors of color class recognition using CVS

Data model Classifier	Image analysis					
	Training errors1,%		Testing errors1,%		Testing errors2,%	
Class	g_i	e_i	g_i	e_i	g_i	e_i
1cc	0	0.8	8.9	12.8	8.9	8
2cc	3	0.2	7.2	4.9	8.3	4.9
3cc	0	0	0	15.9	0	15.9
5cc	0	0	8.3	8.3	79.4	8.3
7cc	2.3	1.4	4.5	13.5	35	13.5
8cc	0	0	0	5.1	0	5.3
9cc					23.2	68.4
$e_o, \%$		0.7		10		30.6

The classification errors are calculated using the equations:

$$e_i = \frac{FP_i}{(TP_i + FP_i)} \quad (1)$$

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

$$g_i = \frac{FN_i}{(TP_i + FN_i)} \quad (2)$$

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 the i -th class classified incorrectly to other classes;

$$e_o = \frac{\sum_{i=1}^N FP_i}{\left(\sum_{i=1}^N TP_i + \sum_{i=1}^N FP_i \right)} \quad (3)$$

e_o (classification error rate) gives the relative part of all incorrectly classified objects, where N is the number of classes.

Analysis of the results

1. The testing results concerning object color zones extraction and color class recognition are acceptable, bearing in mind the nature of the investigated objects. The error $e_o=10\%$ is obtained without elements from class 9cc in testing set. When we include such kind of elements (168 in number), the classification error increases to 41.9%. This result is due to the fact, that in a big part of non-grain impurities color zones typical for the different grains are found.

2. The selection of a proper classifier for color class recognition has a significant influence over the classification accuracy for training and testing. In color class recognition using CDRBE, training and testing errors decrease with 6.2% and 35.1% comparing to CSRBE [M+11]. When we train and test classifiers, CRBEP do not show its advantages because during these two procedures the data for different classes is being processed consecutively. When classifying unknown objects the input vectors will be processed in a random way so it is expected for CRBEP to perform better than the other two.

2.3. Objects shape analysis using CVS

The results from objects shape recognition using the selected classifier are presented in Table 4 (non-grain impurities are not included in the training set).

Table 4. Classification errors of objects shape recognition using the selected classifier

Selected classifier-CRBEP				
Validation error=24.3%				
Train. errors1,%		Test. errors,%		
Class	g_i	e_i	g_i	e_i
1csh	0	0	35.2	4.9
2csh	0	0	60	71.4
3csh	0	0	92.9	72.7
4csh			15.8	39.8
$e_o,%$	0		35.6	

Analysis of the results

1. The results obtained from the test concerning the classification of the investigated objects by its shape show that the rate of objects from class 1csh assigned to other classes is comparatively small (4.9%). On the other hand, the rate of objects assigned to this class, which actually belong to other classes, is sufficiently bigger (35.2%). The rate of objects from class 3csh assigned to 2csh and 4csh is big too.

2. The classification error rate of objects from class 3csh (parts of kernels) is large. This is an expected result because it is impossible to define some standard shape for objects from this class. In many cases, even a qualified expert will not recognize such objects if no color characteristics but only shape is taken into consideration. During the classifiers training models of broken kernels are created on the basis of whole kernel models and that is why the training sample classification error rates for classes 2csh and 3csh are small. This is the explanation of the big difference between training and testing classification results for these two classes.

3.4. Classification in color classes using spectra analysis

The classification results for sample elements color class recognition using the selected classifier (CDRBE) and the three data models are presented in Fig. 5.

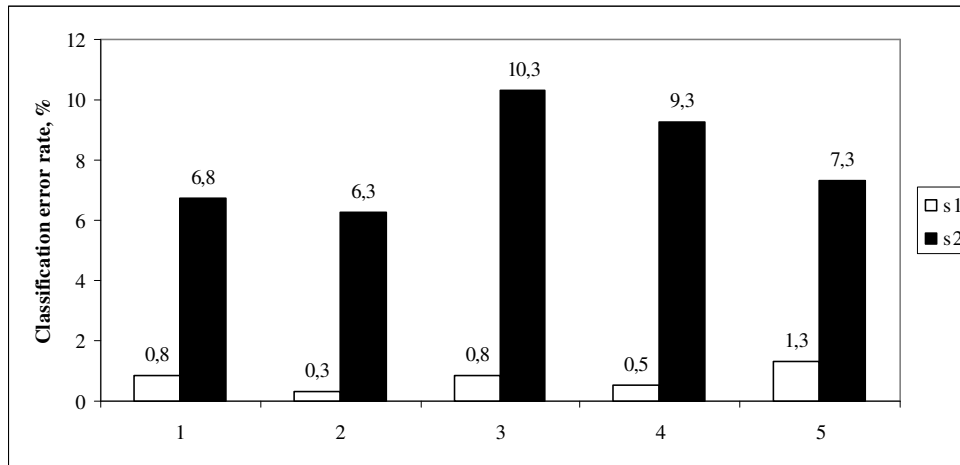


Fig. 5. Classification results obtained using spectra analysis: 1,2,3-training errors: 1-PCA model, 2-Wavelet1+PCA model, 3-Wavelet2+PCA model; 4-validation errors; 5-testing errors: s1-non grain impurities are excluded from validation and testing sets, s2-non grain impurities are included in validation and testing sets

Analysis of the results

1. The testing error obtained (7.3%) is acceptable bearing in mind the specific investigation conditions and the diversity of grain sample elements.

2. The comparative analysis of the results obtained using different variants of classifier validation, training and testing confirms the effectiveness of the INTECHN classification strategy, classifiers, validation approach and data models. For example, if we use the three data models: PCA, Wavelet1+PCA and Wavelet2+PCA the training errors are 6.8%, 6.3% and 10.3% respectively using the CDRBE classifier. The INTECHN validation approach (when the non – grain impurities are included in validation procedure, but are excluded from training sets) decreases the testing error 3.8 times (from 27.6% to 7.3%) in comparison with the traditional validation approach (when the non – grain impurities are simultaneously excluded or included in validation and training sets). The choice of an appropriate classifier for specific classification task has an influence over the classification accuracy too. For example, the training errors obtained using the CDRBE, CSRBE and CRBEP classifiers and PCA data model are 6.8%, 72%, and 7.3% respectively.

3. The classification errors for color class recognition using spectra analysis are sufficiently smaller than the errors using image analysis. For example, the testing errors are 1.3% and 10% respectively using the two approaches when the non – grain impurities are excluded from the validation and testing sets. When we include the non – grain impurities in validation and testing

sets these two errors are 7.3% and 42%. The big difference between the two errors can be explained by the fact that the object spectral characteristics contain not only information for objects color characteristics, but for their surface texture too. Although typical for some grain groups color zones are found in a big part of non – grain impurities, the surface texture of these elements is sufficiently different from the typical for the grains.

2.5. Classification in normative classes

The results from objects classification (when non – grain impurities are included in testing sets) in normative classes when the selected classifiers and the three variants of data fusion are used, are presented in Table 5.

Table 5. Classification errors in normative classes

Color and shape data fusion						
Fusion variant	Variant 1		Variant 2		Variant 3	
Selected data model	RGB		RGB		Wavelet1+PCA	
Selected classifiers	CDRBE– CRBEP		CDRBE– CRBEP		CRBEP-CDRBE	
Errors	Test. errors2,%		Test. errors2,%		Test. errors2,%	
Class	g_i	e_i	g_i	e_i	g_i	e_i
1cst	7.7	28.2	4.2	1.7	2.7	6.8
2cst	27.9	32.2	17.2	14.4	1	11.9
3cst	12.7	0.8	6.4	9.1	8.4	0.4
$e_o, %$	15.3		8.6		5.3	

A detailed representation (confusion matrices) of the grain sample testing set classification using the selected classifier and data model is shown in Tables 6 to 8.

Table 6. Testing set classification (shape recognition) using the selected classifier

		Actual classes			
		1	2	3	4
Predicted classes	1	116	11	0	61
	2	0	18	4	23
	3	3	9	3	27
	4	3	25	4	170
Total number		122	63	11	281

Table 7. Testing set classification (color class recognition) using the selected data model and classifier

Data model: Wavelet1+PCA, classifier: CDRBE										
		Actual classes								
		1	2	3	4	5	6	7	8	9
Predicted classes	1	25	0	0	0	0	0	0	0	0
	2	0	28	0	0	0	0	0	0	0
	3	1	0	17	0	1	0	0	0	0
	4	0	0	0	0	0	0	0	0	0
	5	0	0	0	0	9	0	0	0	0
	6	0	0	0	0	0	0	0	0	0
	7	1	2	0	0	0	0	46	0	3
	8	0	0	0	0	0	0	0	9	0
	9	3	0	4	0	1	0	2	2	131
Total number		30	30	20		13		48	11	134

Table 8. Testing set classification (normative class recognition)

Data fusion – Variant 1				
		Actual classes		
		1	2	3
Predicted classes	1	84	5	2
	2	31	80	0
	3	2	33	240
Total number		117	118	242
Data fusion – Variant 2				
		Actual classes		
		1	2	3
Predicted classes	1	115	3	2
	2	1	101	20
	3	1	14	220
Total number		117	118	242
Data fusion – Variant 3				
		Actual classes		
		1	2	3
Predicted classes	1	109	3	0
	2	0	101	1
	3	8	14	241
Total number		117	118	242

Analysis of the results

1. The object classification in normative classes (1cst, 2cst, and 3cst) involves complex assessment of color and shape characteristics of the investigated objects. Color and shape data are fused together for this purpose. The data fusion procedure improves sufficiently the final classification results. The classification error rate e_0 in normative classes using CVS (Selected variant CDRBE– CRBEP) is 15.3% when data fusion Variant 1 is used and 8.6% when Variant 2 is used, while the errors of object color zones extraction and object shape recognition are 41,9% and 35.6% respectively. .

2. The choice of an appropriate procedure for fusion the results from color characteristics and objects shape analysis has a significant influence over the final classification accuracy. When we use Variant 3 for class recognition, the classification error rate e_0 decreases 1.6 times in comparison with the best results obtained using CVS. This is because the spectra analysis gives the best results for color class recognition.

Conclusions

The results from the investigation at this stage of the INTECHN project implementation concerning estimation of grain sample quality using complex assessment based on color image and spectra analysis can be summarized as follow:

1. The developed approaches, methods, and tools for grain samples quality assessment based on the complex analysis of object color, surface texture, and shape give an acceptable accuracy under specific experimental circumstances.

The error rate $e_0=5.3%$ of the final categorization in the normative classes can be accepted as a good result at this stage of project implementation.

2. The data fusion procedure improves sufficiently the final classification results. The classification error rate e_0 using CVS is 15.3% when Variant 1 is used and 8.6% when Variant 2 is used, while the errors of object color zones extraction and object shape recognition are 41,9% and 35.6% respectively.

3. The results obtained show that the choice of an appropriate procedure for fusion the results from color characteristics and objects shape analysis has a significant influence over the final classification accuracy.

When we use the second algorithm (Variant 2), which is based on color zones topology assessment the classification error rate decreases 1.8 times compared to the first algorithm (Variant 1) in which color class assessment is based on the registration the presence of the typical color zones combinations only. When we fuse the results from color class recognition obtained using spectra analysis and shape class recognition obtained using image analysis (Variant 3) the final classification accuracy is increased 2.9 and 1.6 times in comparison with Variant 1 and Variant 2 respectively.

Acknowledgements

INTECHN platform development, as well as presented analyses and results are a part of implementation of the research project “Intelligent Technologies for Assessment of Quality and Safety of Food Agricultural Products”, funded by the Bulgarian National Science Fund.

References

- [BPS06] **T. Baye, T. Pearson, A. Settles** – *Development of a calibration to predict maize seed composition using single kernel near infrared spectroscopy*, Journal of Cereal Science, vol. 43, No. 2, pp. 236-243, 2006.
- [BS03] **T. Brosnan, D.W. Sun** – *Inspection and grading of agricultural and food products by computer vision systems - a review*, Computers and Electronics in Agriculture, vol. 36, pp. 193 – 213, 2003.
- [C+06] **D. Cozzolino, A. Fassio, E. Fernández, E. Restaino, A. La Manna** – *Measurement of chemical composition in wet whole maize silage by visible and near infrared reflectance spectroscopy*. Animal Feed Science and Technology, Vol. 129, No 3-4, pp. 329-336, 2006.
- [Car03] **M. Carlos** – *Prediction Chemical Composition and Alveograph Parameters on Wheat by Near-Infrared Transmittance Spectroscopy*. Journal of Cereal Chemistry, Vol. 51, No 21, pp. 6335–6339, 2003.

- [F+09] **V. Fernández-Ibañez, A. Soldado, A. Martínez-Fernández, B. de la Roza-Delgado** – *Application of near infrared spectroscopy for rapid detection of aflatoxin B1 in maize and barley as analytical quality assessment*. Food Chemistry, Vol. 113, No 2, pp. 629-634, 2009.
- [H+08] **H. Huang, H. Yu, H. Xu, Y. Ying** – *Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review*, Journal of Food Engineering, vol. 87, No. 3, pp. 303-313, 2008.
- [JPV00] **D. Jayas, J. Paliwal, N. Visen** – *Multy-layer Neural Networks for Image Analysis of Agricultural Products*. Agricultural Engineering Research, Vol. 77, No 2, pp. 119-128. 2000.
- [L+93] **K. Liao, M. Paulsen, J. Reid, B. Ni, E. Maghirang** – *Corn kernel breakage classification by machine vision using a neural network classifier*, Transactions of the ASAE, vol. 36, No. 6, pp. 1949 – 1953, 1993.
- [LJS99] **X. Luo, D.S. Jayas, S. Symons** – *Comparison of statistical and neural network methods for classifying cereal grains using machine vision*. Transactions of the ASAE, Vol. 42, No. 2, pp. 413 – 419, 1999.
- [M+02] **M. Mladenov, S. Penchev, R. Tzonev, P. Daskalov, V. Stoianov, E. Kuleva, B. Borisov** – *Seeds Purity Assessment Using 2D Image Analysis*. In Proceedings of the 4th International Conference on Technology and Automation, Thessaloniki, Greece, pp. 399 – 404, 2002.
- [M+04] **M. Mladenov, S. Penchev, B. Borisov, K. Arvanitis, N. Sigrimis** – *Evaluation of some properties for purity assessment of seeds using computer vision system*, In Proceedings of AgEng2004, Leuven, Belgium, paper No. 590, 2004.
- [M+10] **S. Mahesh, D. Jayas, J. Paliwal, N. White** – *Identification of wheat classes at different moisture levels using near-infrared hyperspectral images of bulk samples*, Sensing and

instrumentation for food quality and safety, DOI: 10.1007/s11694-010-9104-2, 2010.

- [M+11] **M. Mladenov, S. Penchev, M. Dejanov, M. Mustafa** – *Quality Assessment of Grain Samples using Color Image Analysis*, In Proceedings of the IASTED International Conference, Signal Processing, Pattern Recognition, and Applications (SPPRA 2011), Innsbruck, Austria, pp. 37 – 44, 2011.
- [MD08] **M. Mladenov, M. Dejanov** – *Application of neural networks for seed germination assessment*, In Proceedings of the 9th WSEAS international conference on Neural Networks CNN'08, Sofia, Bulgaria, pp. 67 – 72, 2008.
- [MJ00] **S. Majumdar, D. Jayas** – *Classification of cereal grains using machine vision, Part3: Texture models*, Transactions of the ASAE, vol. 43, No. 6, pp. 1681 – 1687, 2000.
- [N+88] **S. Ning, R. Ruan, L. Luo, X. Chen, P. Chen, R. Jones** – *Automation of a machine vision and neural network based system for determination of scabby rate of wheat samples*, ASAE Annual International Meeting, Orlando, Florida, USA, 1988.
- [O+07] **J.I. Ortiz-Monasterio, N. Palacios-Rojas, E. Meng, K. Pixley, R. Trethowan, R.J. Peña** – *Enhancing the mineral and vitamin content of wheat and maize through plant breeding*. Journal of Cereal Science, Vol. 46, No 3, pp. 293-307, 2007.
- [P+03] **J. Paliwal, N. Visen, D. Jayas, N. White** – *Cereal grain and dockage identification using machine vision*, Biosystems Engineering, vol. 85, No. 1, pp. 51-57, 2003.
- [P+10] **K.H.S. Peiris, M.O. Pumphrey, Y. Dong, E.B. Maghirang, W. Berzonsky, F. E. Dowell** – *Near-Infrared Spectroscopic Method for Identification of Fusarium Head Blight Damage and Prediction of Deoxynivalenol in Single Wheat Kernels*. Journal of Cereal Chemistry, Vol. 87, No 6, pp. 511–517, 2010.

- [S+05] **M. Singh, M.R. Paulsen, L. Tian, H. Yao** – *Site-specific study of corn protein, oil, and extractable starch variability using NIR spectroscopy*. Applied Engineering in Agriculture, Vol. 21, No 2, pp. 239-251, 2005.
- [W+04] **Z. Wang, J.H. Wang, L.Y. Liu, W.J. Huang, C.J. Zhao, C.Z. Wang** – *Prediction of grain protein content in winter wheat (*Triticum aestivum* L.) using plant pigment ratio (PPR)*. Field Crops Research, Vol. 90, No 2-3, pp. 311-321, 2004.

TIBISCUS