U.P.B. Sci. Bull., Series C, Vol. 73, Iss. 4, 2011 ISSN 1454-234x

# AUTOMATIC CLASSIFICATION OF GRAIN SAMPLE ELEMENTS BASED ON COLOR AND SHAPE PROPERTIES

Miroljub MLADENOV[1], Stanislav PENCHEV[2], Martin DEJANOV[3], Metin MUSTAFA[4]

*An approach for assessment of the main quality features of grain sample elements using image analysis is presented. The goal is to classify the sample elements to one of the normative groups, depending on the object color and shape features. The assessment of these features is performed separately, whereupon the results from the two assessments are fused to make the final object's classification. Methods and tools for object color and shape analysis and assessment, as well as for fusing the results from these two assessments are proposed. The performance of the procedures developed is evaluated.*

**Keywords:** grain quality assessment, computer vision, feature extraction, classification

## 1. Introduction

The main characteristics of the grain quality are grain appearance, shape, color, smell, flavour, moisture content, presence of impurities – grain and non – grain ones. The whole grains with appearance, shape and color inherent for the variety and hybrid, as well as the broken grains bigger than the half of the whole grain, are considered standard. There are several groups, which are considered as grain impurities: broken grains smaller than the half of the whole grain, heat-damaged grains, small, shriveled and green grains, sprouted grains, moldy grains and infected grains. The group of the non – grain impurities consists of: corn-cob particles, leaf and stem fractions, pebbles, soil, sand, dust and metal particles, smutty grains, as well as harmful elements (bunt).

Most of the grain quality characteristics mentioned above (except moisture content, smell and flavour) are related to the color, shape and dimensions of the grain sample elements. The main trend in the last few years is to use Computer

[1] Assoc. Professor, PhD, Department of Automatics, Information and Control Engineering, University of Ruse, Bulgaria, e-mail: mladenov@uni-ruse.bg
[2] Principal Assistant, Department of Automatics, Information and Control Engineering, University of Ruse, Bulgaria, e-mail: msp@uni-ruse.bg
[3] Principal Assistant, PhD, Department of Automatics, Information and Control Engineering, University of Ruse, Bulgaria, e-mail: mdejanov@uni-ruse.bg
[4] PhD student, Department of Automatics, Information and Control Engineering, University of Ruse, Bulgaria, e-mail: mmustafa@uni-ruse.bg

Vision Systems (CVS) for evaluation of such characteristics [1]. Many results are published, in which color and color texture analyses were used to assess some particular quality features. In [2] 25 textural features of individual kernels are used in the assessment of the authenticity of five different kinds of cereal grains. The textural features are extracted from different colors and color band combinations of images. Different grain color features are used to assess variety [3], infections [4], germination [5], weed identification [6], etc.

Morphological features, related to the grain shape and geometrical parameters are used for assessment of the grain variety too. A set of eight morphological features is presented in [7] to recognize five different kinds of cereal grains. A broader investigation, with a total of 230 features used for classification of barley, Canada Western Amber Durum wheat, Canada Western Red Spring wheat, oats, and rye is presented in [8]. Assessment of the grain sample purity is performed by profile analysis of corn kernels using one-dimensional digital signals based on its binary images [9], by modeling the shape using a set of morphological features [10] and by shape curvature analysis [11]. Mechanical damage was determined using both single-kernel and batch analysis by extracting from kernel images the damaged area stained by green dye and by calculating the percentage of total projected kernel surface area that was stained green [12]. Mold damage was determined using single-kernel analysis by isolating the moldy area on kernel images and by calculating the percentage of total projected kernel surface area covered by mold [12].

The references cited above present some results concerning the assessment of specific grain quality features. The main goal of this paper is to present an approach for a complex assessment of the maize grain quality. This approach is based on the analysis of the object color, shape and dimensions using CVS, as well as the fusion of the results of the separate analyses. The categorization of the grain sample elements in three quality groups is the final goal of this complex assessment. The approach is realized within the frame of the INTECHN project "Development of Intelligent Technologies for Assessment of Quality and Safety of Food Agricultural Products", funded by the Bulgarian National Science Fund. The INTECHN platform methods and tools for recognition of the color characteristics and shape of grain sample elements, as well as the methods for combining the results from color and shape analysis, are presented. Some of the results obtained at this stage of project implementation are given too.

## 2. Materials and methods

### 2.1. Image acquisition hardware.

Within the frame of the INTECHN project, CVS is used (Fig. 1) to assess such characteristics of grain sample elements, which are usually evaluated by an expert based on visual estimation only. The images of the grain sample elements

(2) are taken using a color CCD camera, model DFK31AU03 (THE IMAGINGSOURCE, Germany) (1) with lens PENTAX B2514D (Hoya Corporation, Japan) (3). The illumination system includes two luminescent ring-shaped units with different diameters (4). A portable computer model Dell Vostro 1720 is used for implementation of the INTECHN platform software procedures.

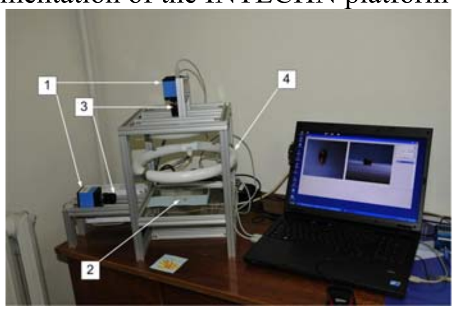

Fig. 1. INTECHN platform CVS

### 2.2. Grain sample classification groups

Groups (classes) and subgroups (subclasses), in which the elements from the maize grain sample have to be distributed, are presented in Table 1.

*Table 1*

**Maize grain 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<br>2cc- grains with color inherent for the variety, germ side | 1csh- whole grains with inherent for the variety shape |
| 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 *Fusarium*) grains, smutty grains. | 3cc- heat-damaged grains<br>4cc- green grains<br>5cc- moldy grains<br>6cc- smutty grains | 2csh- broken grains bigger than the half of the whole grain<br>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 *Fusarium*) grains<br>8cc- sprouted grains<br>9cc- non grain impurities | 4csh- non grain impurities |

It is evident from the table, that the main characteristics, which determine the grouping of grain sample elements in classes and subclasses, are related to the objects' color and shape. Because of the color and shape features are described in a different manner, it is advisable the assessment of these characteristics to be made separately. After that, the results from the two assessments have to be fused to obtain the result about the object's classification in one of the normative classes.

Maize grain samples of the Kneja-433 variety are used in the investigation. The Maize Research Institute-Kneja, Bulgaria, produces this variety. The samples are gathered in one corn-growing season of one crop year and from one growing location.

### 2.3. Determination of the color class of grain sample elements

The analysis of the color characteristics of maize grain sample elements shows, that the following typical color zones can be found in different sample element images: color zone 1cz–grain back side zone, 2cz–grain germ zone, 3cz–heat-damaged grain zone, 4cz–green grain zone, 5cz–mold grain zone, 6cz–smutty grain zone, 7cz–*Fusarium* infected grain zone, 8cz–germ zone (in sprouted grains), 9cz- grain tip cap zone, 10cz-grain crown zone and 11cz. The 11cz color zone includes pixels, whose color characteristics are different from the pixel features of the other ten color zones. Each of the zones defined consists of a set of neighboring pixels, which have similar color features.

*Table 2*

**Color class determination using typical combinations of color zones**

| **Color zones** | | | | | | | | | | | **Color class** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0/1 | 1 |
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0/1 | 2 |
| 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0/1 | 3 |
| 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0/1 | 4 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0/1 | 5 |
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0/1 | 7 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0/1 | 8 |

The decision if the object corresponds to one of the color classes presented in Table 1, depends on the presence in the object region of a typical for this class combination of color zones. The possible combinations for each of the color classes are predetermined. Examples of color zone combinations, which correspond to different color classes, are presented in Table 2. Therefore, the task for recognition of the color class is reduced to a task for detecting a typical combination of color zones within the object region.

#### 2.3.1. Extraction of color zones from the object image

The extraction of a specific zone in a color image is a typical classification task for association of the pixels to one of the predetermined classes. This procedure includes the following steps: development of color zone models, evaluation of the color features of each pixel from the object region, followed by the association of the pixels to one of the models defined and detection of regions with neighboring pixels, whose number exceeds some threshold value.

**Development of color zone models.** The following approach is used to develop color zone models. Firstly, the operator marks consecutively regions in the grain image, which include pixels with features inherent for the specific color zone. This action is repeated in different grain images. In this way, a set of pixels (training set) with similar color features is completed for each color zone. These sets are used to develop models of the first ten color zones. If we present these sets of pixels in the RGB space, it can be seen that they form comparatively compact regions.

The color zone models are represented by the zone centre (the average of the RGB values of the pixels from the training set) and the zone boundary surface. The boundary surface is determined through a threshold value of the covariance of the RGB values. It is relevant to remark that models are created only for the first ten zones, presented in Table 2. A correct model for 11cz color zone cannot be created because the color features of the pixels from this zone are sufficiently different in each subsequent grain sample.

The task for color zone modeling is reduced to a task for approximation of the color zone regions. For this purpose specific classifier architectures, based on Radial Basis Elements (RBEs) are developed. Their application is determined by 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 $\Delta$ of its output, it becomes clear what part of input objects will be included within the class boundary and it is easy to change the dimensions of the particular class region.

The following classifiers [13] are used for class region approximation:

**Classifier with standard RBEs (CSRBE).** Only one RBE is used for approximation of each class area. 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 x 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.

#### 2.3.2. Association of the object pixels with one of the color zones defined

**Description of the pixel color features.** The possibility for using the RGB, HSV and XYZ color models for description of the pixel color features is analysed. Four color texture models [13] are developed for the same purpose. It is expected that these models will better underline color ratios of the pixels from different color zones. The texture models can be presented by the following equations:

**First texture model**. It is constructed on the basis of RGB model. Its components are the normalized differences between the R, G and B components:

$$d_1 = \frac{R-G}{R+G+B}\ ;\ d_2 = \frac{B-R}{R+G+B};\ d_3 = \frac{G-B}{R+G+B};\ d_1 + d_2 + d_3 = 0 \qquad (1)$$

**Second texture model**. This model includes non-linear transformation onto R, G and B components as follows:

$$O_1 = \frac{R}{G}; O_2 = \frac{B}{R}; O_3 = \frac{G}{B};\ \mathrm{O_1 O_2 O_3} = 1 \qquad (2)$$

**Third texture model**. It is similar to the second model. The pixel intensity is added as fourth coordinate:

$$O_1 = \frac{R}{G}; O_2 = \frac{B}{R}; O_3 = \frac{G}{B};\ I = R + G + B; \mathrm{O_1 O_2 O_3} = 1 \qquad (3)$$

**Fourth texture model.** This model converts input RGB space into one-dimensional texture feature $T_k$:

$$T_k = \frac{3(R + mG + nB)}{(R + G + B)(1 + m + n)}; m = \frac{G}{R};\ n = \frac{B}{R} \qquad (4)$$

For separation of the object region from the background, the best results are obtained using the second texture model. When extracting different color zones within the object region, the best performance have the HSV, RGB color models, as well as the first and second texture model. Therefore, the operator, who performs the INTECHN platform training, has the possibility to choose the appropriate color or texture model based on the results obtained during the training procedure.

**Association of the pixels from the object region with one of the models defined.** As it was mentioned above, in the object image, besides pixels, corresponding to one of the color zones, there are those that have sufficiently different color features. For part of the grain sample elements, for example the non – grain impurities, pixels could get into the boundary of one of the color zone models, but a big part of them would get outside the color zone boundaries or could be located in a random place in the feature space. These pixels could be considered as noisy vectors. It could be assumed that the comparatively compact color zone regions are submerged in a noisy environment. Therefore, the task for color zone extraction could be interpreted as a task for classification in classes, whose boundaries have definite shapes, dimensions and location in the feature space, but are situated in a noisy environment.

Having in mind this formulation, the use of popular classification strategies like discriminant analysis, cluster analysis, support vector machines, K–nearest neighbors and some others, which build boundaries between class regions, is obviously not a good choice. This is due to the fact, that for the class, which corresponds to the color zone 11, a correct model cannot be created. If we develop a model of this class using training set consisting of non-grain impurities, it could be expected that a big part of the pixels of elements from the testing set of this group would get outside the class model. Thus, this will be an incorrect classification.

In unison with this formulation, models of the classes, corresponding to the first ten zones, are created using pixels from the respective training sets. The description of each pixel from the object region is used as a classifier input. The output with the maximum value is selected. If this value exceeds the threshold $\Delta$, the pixel is associated with the respective color zone model. Otherwise, it is rejected by the classifier and is considered as a noise.

**Detection of color zones with an excess number of pixels.** To detect a color zone, a set of pixels with color features inherent for this zone has to be detected within the frame of the object region. The number of pixels has to exceed some threshold value. The classifier for color zone recognition (CCZR) is presented in Fig. 2. It is used for detection of different color zones in the object region.

The classifier architecture includes ten analogous parallel channels. Each channel is used to separate pixels from a particular color zone. The first classifier layer transforms the RGB pixel description into the preliminary chosen color or texture model. The second layer includes structures for modeling of each typical color zone (CSRBE or CDRBE). The next layer includes ten accumulative elements, which are used to count the pixels, classified in each of the color zones.

The last classifier layer consists of threshold elements for each color zone. They determine whether the accumulated number of pixels of the respective zone exceeds the threshold value of the ratio between those accumulated pixels and the total number of pixels in the object region. The threshold value for each of the zones is predetermined after an appropriate analysis during the training procedure. While the previous layers are used for the analysis of each pixel in the image, the threshold layer is used after the analysis of all the pixels in the object region.

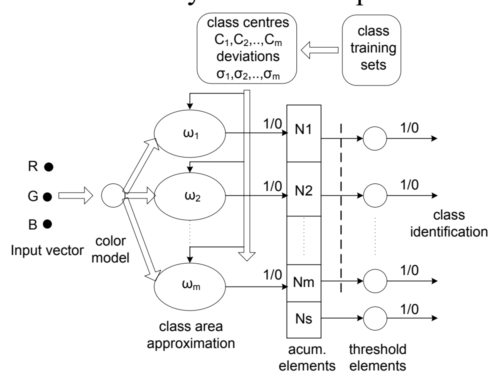


Fig. 2. Classifier for detection of grain color zones (CCZR)

### 2.4. Determination of the grain sample elements' shape

#### 2.4.1. Classification groups related to the shape of the grain sample elements

In keeping with Table 1, according to the shape, the grain sample elements are divided into three main classes (1csh, 2csh and 3csh) and one additional class (4csh), which corresponds to non-grain impurities. To increase the accuracy of the shape recognition procedure, each of the main classes is additionally divided into six subclasses. This is made to obtain comparatively compact class regions. Such region can not be created for the 4csh class, because the shape of the elements from this group varies within a wide range.

#### 2.4.2. Shape description of the grain sample elements

To represent the shape of the grain sample elements ten – dimensional vector descriptions are used. The following approach is used to obtain shape description of the investigated objects. First, the binary image of the object zone is created. After that the object's peripheral contour is extracted, the bisection line of the contour is detected and an odd number of cross – sections perpendicular to the bisection line are built (Fig. 3).

The cross – sections relative length $h_i = s_i/D$, as well as the size and the sign of the difference between two neighbour sections $\Delta_i = \Delta_1 - \Delta_2$ are calculated. Finally the object shape description is presented in the following form:

$$X_{sh} = \left(h_1, h_2, \ldots, h_n, \Delta_1, \Delta_2, \ldots, \Delta_n\right) \tag{5}$$

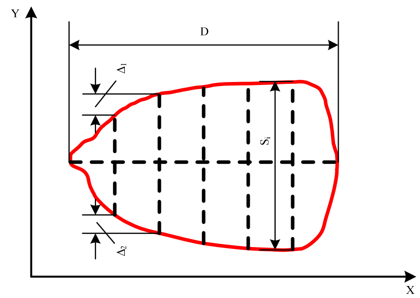


Fig. 3. Object shape description: D – length of the bisection line; $h_i = s_i/D$ – length of a cross – section; $\Delta_i = \Delta_1 - \Delta_2$ - difference between neighbor cross – sections.

It is typical for maize kernels that their contour line has a huge asymmetry at the part where the grain tip cap exists and at the opposite part (grain crown region). It is easy to locate the tip cap region in the whole grains and to build contour descriptions and models with proper orientation. For broken grains, depending on what part of the whole grain has remained (the one with the germ or the one without it) the contour descriptions can be sufficiently different. That is why it is necessary for 2csh and 3csh classes (or, more precisely, for their corresponding subclasses) to define two types of descriptions and models: for shapes where the tip cap exists in the remaining part of the grain, and for those without a tip cap.

**2.4.3. Development of shape class models of the grain sample elements**

The task for object shape recognition is divided into two stages: 1) development of shape subclass models and 2) association of an unknown object to one of these subclasses.

The task for development of shape subclass models is additionally divided into two steps. The first step includes the creation of the training set of contour descriptions for each of the six subclasses that correspond to the whole grains. The average description and the covariance are calculated from the training set of each group. The average description is used as a centre of a particular subclass, while the covariance defines its boundary.

At the second step, artificial contour models of the broken kernels are generated from the contour models of the whole grains. It is done by setting to zero the last, the last two, the last three and the last four cross – sections from the respective whole grain description. Such approach is conditioned by the fact that it is impossible to create a compact model of real broken grains, because the shape of these elements varies within a wide range.

Like the task for development of color zone models, the task for development of shape class models is reduced to a task for approximation of the

18 shape subclasses. The possibility for using the CSRBE and CDRBE classifiers is investigated for this purpose. Furthermore, another variant of a classifier based on RBEs is developed. It is intended for classification in overlapping classes, as some of the shape classes are expected to be so.

**Classifier with RBEs which takes into consideration the class potentials (CRBEP).** The CRBEP classifier [13] approximates the class areas using standard (or decomposing) RBEs and takes into consideration the class potentials. The accumulated during classification number of vectors in each of the classes 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.

#### 2.4.4. Association of an unknown object to one of the shape class models

The shape description of each unknown object is used as a classifier input. The output with the maximum value is selected. If this value exceeds the threshold $\Delta$, the unknown object is associated with the respective shape subclass model. Otherwise, the classifier rejects it. If the input object is not associated with one of the shape subclasses, it is classified to the fourth class.

### 2.5. Fusing of the results from the color and shape analysis

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 simple fusion scheme is used in the first algorithm, presented in Fig. 4. The input data (input classes) is separated in two groups – objects' color data and objects' shape data. The first group contains the ten color zones, described in section 2.3. The second group consists of the eighteen shape subclasses.

The color class (1cc, 2cc… 8cc) is determined at the first stage of data fusion based on predetermined combinations of color zones. The shape subclasses are merged into the three main shape classes (1csh, 2csh and 3csh).

The color and shape classes are fused at the second stage of the algorithm.

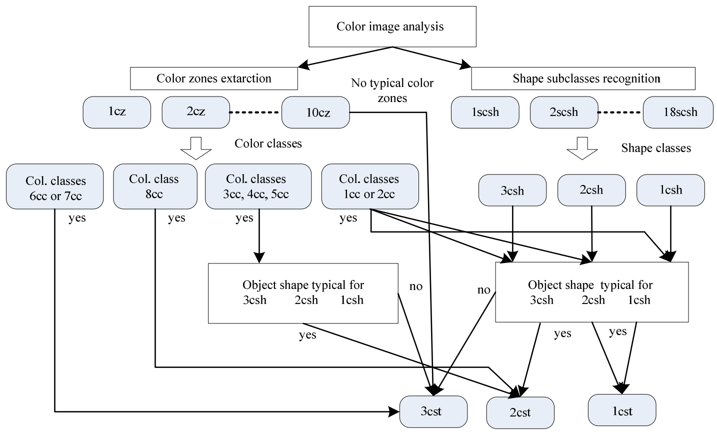


Fig. 4. Fusing of the results from color and shape analysis. Variant 1.

The result of this procedure is the final decision for object classification in one of the three normative classes (1cst, 2cst and 3cst). The assessment whether the object shape is classified in one of the three main shape classes or not is used as a fusion criterion for the color classes 1cc to 5cc. The shape is not important for the 6cc, 7cc and 8cc classes.

**Variant 2.** Color topological and combined topological models of color zones are used in the second algorithm (Fig. 5).

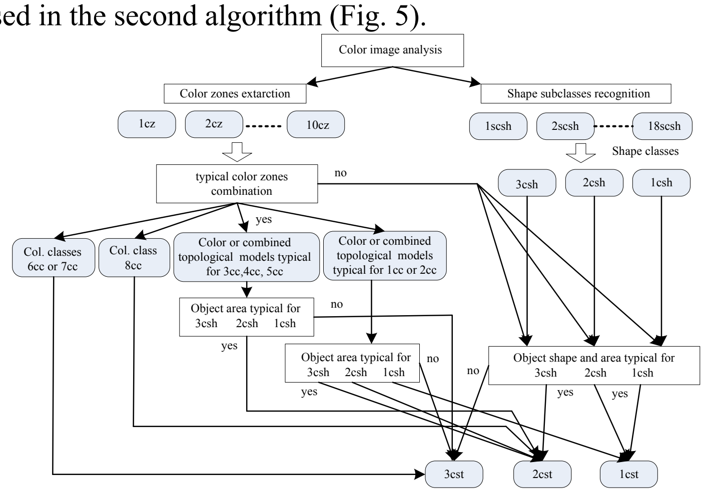


Fig. 5. Fusing of the results from color and shape analysis. Variant 2.

The topological models represent the plane distribution of the color zones within the object region. Such models are created when three or more color zones typical for the grains are found in the object region. The combined topological models represent the plane distribution of some shape elements (the region of grain tip cap or crown region) and the color zones found. These models are created when only two typical color zones are found.

The final categorization of the objects, for which such topological models are found, is performed taking into account the object's area as a criterion. When only one typical color zone is found, the object shape and the object area are the most important for the final categorization. When the object belongs to the 6cc, 7cc or 8cc classes, the object shape is not important for its categorization.

## 3. Experimental results

### 3.1. Training and testing sets

The developed procedures for grain quality assessment are validated, trained, and tested with the sets, presented in Table 3. Elements from the 4cc and 6cc color classes are not included in color class recognition because such elements are rarely met in maize grain samples.

*Table 3.*

**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 | 256 | | | |
| **Classification in normative classes** | | | | | | | |
| Classes | 1cst | 2cst | 3cst | | | | |
| Testing sets | 117 | 118 | 242 | | | | |

### 3.2. Results from the object color analysis

The classification errors in color class recognition are presented in Table 4. The Testing error1 represents the results when the group of the non-grain impurities is not used for testing. The Testing error2 represents the case, when this group is included in the testing samples.

*Table 4*

**Classification errors in color class recognition**

| CSRBE | | Selected classifier – CDRBE | | |
|---|---|---|---|---|
| Train. error | Test. errors | Train. errors | Test. errors1 | Test. errors2 |
| $e_o$=6.9% | $e_o$=77% | $e_o$=0.7% | $e_o$=10% | $e_o$=30.6% |

The classification error rate $e_o$ is calculated using the equation:

$$e_o = \frac{\sum_{i=1}^{N} FP_i}{\left(\sum_{i=1}^{N} TP_i + \sum_{i=1}^{N} FP_i\right)} \tag{6}$$

It gives the relative part of all incorrectly classified objects, where N is the number of classes; $FP_i$ is the number of elements from the i-th class, which are incorrectly classified to other classes and $TP_i$ is the number of elements from the i-th class, which are correctly classified.

**Analysis of the results**

1. The testing results concerning the object color zone 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 the class 9cc in the testing set. When we include such kind of elements (256 in number), the classification error increases to 30.6%. This result is due to the fact, that typical for the grains color zones are found in a big part of the non-grain impurities. The big testing error is conditioned by the big percentage of the non-grain impurities in the testing set too. The percentage of these elements is intentionally increased. It is about 57% of the all elements in the testing set.

2. The selection of a proper classifier for color class recognition has a significant influence over the training and testing classification accuracy. When the CDRBE classifier is used, the training and testing errors decrease with 6.2% and 67% respectively, comparing to the referent CSRBE classifier.

### 3.3. Results from the object shape analysis

The results from the object shape recognition are presented in Table 5 (non-grain impurities are included in the testing sets).

*Table 5*

**Classification errors in object shape recognition using CSRBE, CDRBE and CRBEP**

| CSRBE | | CDRBE | | Selected classifier-CRBEP | |
|---|---|---|---|---|---|
| Train. errors | Test. errors | Train. errors | Test. errors | Train. errors | Test. errors |
| $e_o$=0% | $e_o$=36.5% | $e_o$=3.6% | $e_o$=27.5% | $e_o$=0% | $e_o$=35.6% |

**Analysis of the results**

1. The results from the classification of the objects from the testing sample in the shape classes show, that the rate of the objects from the 1csh class, which are 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 the objects from the 3csh class, which are assigned to the 2csh and 4csh classes, is big too.

2. The error rate in recognition of the shape class 3csh (parts of grains) is large too. This is an expected result, because a typical shape of the objects from this class cannot be defined. In many cases, even a qualified expert would not recognize such objects if only shape would be taken into consideration. During the classifier training, artificial models of the broken grains are created based on the whole grain models. That is why the classification error rates for training samples of 2csh and 3csh classes are small. This explains the big difference between the training and testing classification results for these two classes.

### 3.4. Classification in the normative classes

The results from object classification in the normative classes (non-grain impurities are included in the testing sets), using the two variants of data fusion, are presented in Table 6. The object distribution in the normative classes (confusion matrix) is shown in Table 7.

*Table 6*

**Testing errors in the normative class recognition**

| Color and shape data fusion | | |
|---|---|---|
| CDRBE–CDRBE | CDRBE–CSRBE | Selected variant CDRBE– CRBEP |
| Variant1 | | |
| $e_o$=15.7% | $e_o$= 15.7% | $e_o$=15.3% |
| Variant2 | | |
| $e_o$= 8.6% | $e_o$= 8.8% | $e_o$=8.6% |

*Table 7*

**Classification of the testing sample elements in the normative classes**

| Color and shape data fusion | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Actual classes | | | | | |
| | | Variant1 | | | Variant2 | | |
| | | 1 | 2 | 3 | 1 | 2 | 3 |
| Predicted classes | 1 | **84** | 5 | 2 | **115** | 3 | 2 |
| | 2 | 31 | **80** | 0 | 1 | **101** | 20 |
| | 3 | 2 | 33 | **240** | 1 | 14 | **220** |
| Total number | | 117 | 118 | 242 | 117 | 118 | 242 |

### Analysis of the results

1. Object classification in the normative classes (1cst, 2cst and 3cst) is based on the complex assessment of the color and shape characteristics of the grain sample elements. For this purpose, the results from color and shape analyses are fused. The data fusion procedure improves sufficiently the final classification accuracy. The classification error $e_0$ (selected variant CDRBE– CRBEP) is 15.3% when data fusion Variant1 is used, and 8.6% - when Variant2 is used.

2. In spite of the big errors in color and shape class recognition (30.6% and 35.6% respectively), the error in the final categorization of the grain sample elements decreases sufficiently (8.6%). This result is due to the application of the

fusion procedures. These procedures decrease or ignore the influence of the factors, which determine the big errors in color and shape class recognition. The main factor, which influences on the accuracy of the color class recognition, is the presence of inherent for the grains color zones in a big part of the non-grain impurities. The big variations of the shape of the broken grains are the most significant factor, which determines the big errors in the shape class recognition.

3. The results from object classification in normative classes show, that the choice of an appropriate procedure for fusing the results from color and shape analyses has a significant influence over the final classification accuracy. When using the second algorithm (Variant 2), which is based on the color or combined topology assessment, the classification error rate decreases almost twice comparing to the first algorithm (Variant 1), in which color class assessment is based only on the presence of the typical combinations of color zones.

## 4. Conclusions

The results from the investigation can be summarized as follows:

1. The developed methods and tools for description and analysis of the color characteristics and the shape of the grain sample elements give the following accuracy at this stage of the project implementation. The testing classification accuracy in recognition of object color classes is 90% and in object shape recognition - 73%, when the non-grain impurities are excluded from the testing sets. The classification accuracy decreases when they are included in the test samples,. The testing errors in recognition of color and shape classes increase with 20.6% and 8.4%, respectively.

2. In spite of the big errors in color and shape class recognition (30.6% and 35.6% respectively), the error in the final categorization of the grain sample elements decreases sufficiently (8.6%). This result is due to the application of the fusion procedures because they decrease or ignore the influence of the factors, which determine the big errors in the color and shape class recognition.

3. The use of the color or combined topology of color zones within the object region leads to an improvement of the final classification accuracy. Under the specific experimental circumstances, the classification accuracy is 91.4% while it is 84.7%, when the color class assessment is based on the combinations of typical color zones only. The classification error rate decreases 1.8 times.

4. The INTECHN validation and training procedures ensure the selection of the optimal model for input data representation, the optimal classifier, and classifier parameters for specific classification task. For example, the selection of a proper classifier for recognition of the object color and shape classes has a significant influence over the classification accuracy. When the CDRBE classifier

is used in color class recognition, the training and testing errors decrease with 6.2% and 67% comparing to the referent CSRBE classifier.

**Acknowledgement**

The INTECHN platform development, as well as the analyses and results presented, are part of the 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

[1] *T. Brosnan, D.W. Sun*. "Inspection and grading of agricultural and food products by computer vision systems - a review". Computers and Electronics in Agriculture, **36**, 2003, pp. 193 – 213

[2] *S. Majumdar, D.S. Jayas*. "Classification of cereal grains using machine vision, Part3: Texture models". Transactions of the ASAE, **43**(6), 2000, pp. 1681 – 1687

[3] *S. Majumdar, D.S. Jayas*. "Classification of cereal grains using machine vision: Color Models", Transactions of ASAE, **43**(6), 2000, pp. 1677 – 1680

[4] *S. Ning, R. Ruan, L. Luo, X. Chen, P.L. Chen, R.K. 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, 1988

[5] *R. Urena, F. Rodrigues, M. Berenguel*. "A machine vision for seeds germination quality evalution using fuzzy logic", Computers and Electronics in Agriculture, **32**(1), 2001, pp. 1 – 43

[6] *M.J. Aitkenhead, I.A. Dalgetty, C.E. Mullins*. "Weed and crop discrimination using image analysis and artificial intelligence methods", Computers and Electronics in Agriculture, **39**, 2003, pp. 157-171

[7] *J. Paliwal, N.S. Visen, D.S. Jayas*. "Evaluation of Neural Network Architectures for cereal grain classification using morphological features", Journal of Agricultural Engineering Research, **79**(4), 2001, pp. 361-370

[8] *J. Paliwal, N.S. Visen, D.S. Jayas*. "Comparison of a neural network and non-parametric classifier for grain kernel identification", Biosystems Engineering, **85**(4), 2003, pp. 405-413

[9] *K. Liao, M.R. Paulsen, J.F. Reid, B.C. Ni, B.E.P. Maghirang*. "Corn kernel breakage classification by machine vision using a neural network classifier", Transactions of the ASAE, **36**(6), 1993, pp. 1949-1953

[10] *J. Paliwal, N.S. Visen, D.S. Jayas, N.D.G. White*. "Cereal grain and dockage identification using machine vision", Biosystems Engineering, **85**(1), 2003, pp. 51-57

[11] *M. Mladenov, S. Penchev, B. Borisov, K. Arvanitis, N. Sigrimis*. "Evaluation of some properties for purity assessment of seeds using computer vision system", Proceedings of AgEng2004, workshop 3, Paper No 590, 2004

[12] *H.F. Ng, W.F. Wilcke*. "Machine vision evaluation of corn kernel mechanical and mold damage". Transactions of the ASAE, **41**(2), 1998, pp. 415-420

[13] *M. Mladenov, S. Penchev, M. Dejanov, M. Mustafa*. "Quality Assessment of Grain Samples using Color Image Analysis". Proc. of the IASTED International Conference SPPRA 2011, February 16 - 18, 2011 Innsbruck, Austria, 2011, pp. 37 – 44.