We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Cognitive and Statistical Pattern Recognition Applied in Color and Texture Segmentation for Natural Scenes

Luciano Cássio Lulio, Mário Luiz Tronco,
Arthur José Vieira Porto,
Carlos Roberto Valêncio and
Rogéria Cristiane Gratão de Souza

Additional information is available at the end of the chapter

1. Introduction

In this approach, cognitive and statistical classifiers were implemented in order to verify the estimated and chosen regions on unstructured environments images. As inspection of crops for natural scenes demands and requires complex analysis of image processing and segmentation algorithms, since these computational methods evaluate and predict environment physical characteristics, such as color elements, complex objects composition, shadows, brightness and inhomogeneous region colors for texture, JSEG segmentation algorithm was approached to segment these ones, and ANN and Bayes recognition models to classify images into predetermined classes (e.g. fruits, plants and general crops). The intended approach to segment classification deploys a customized MLP topology to classify and characterize the segments, which deals with a supervised learning by error correction – propagation of pattern inputs with changes in synaptic weights in a cyclic processing, with accurate recognition as well as easy parameter adjustment, as an enhancement of iRPROP algorithm (*improved resilient back-propagation*) (Igel and Hüsken, 2003) derived from *Back-propagation* algorithm, which has a faster identification mapping process, that verifies what region maps have similar matches through the explored environment. Bayes statistical models had the addiction of process variable as set parameters of predictive error correction.

To carry through this task, a feature vector is necessary for color channels histograms (layers of primary color in a digital image with a counting graph that measures how many pixels are at each level between black and white). After training process, the mean squared error

(MSE), denotes the best results achieved by segment classification to create the image-class map, which represents the segments into distinct feature vectors. Furthermore, a language dictionary is used for the expansion on main results, which semantic regions and negation detection are applied as data mining process with cognitive and statistical classifiers.

2. JSEG image segmentation

Color images with homogeneous regions are segmented with an algorithm to generate clusters in the color space/class (different measures classes in spectral distribution, with distinct intensity of visible electro-magnetic radiation at many discrete wavelengths) (Deng *et* al, 1999a). One way to segment images with textures is to consider the spatial arrangement of pixels using a region-growing technique whereby a homogeneity mode is defined with pixels grouped in the segmented region. Furthermore, in order to segment texture images one must consider different scales of images.

The JSEG algorithm segments images of natural scenes properly, without manual parameter adjustment for each image and simplifies texture and color. Segmentation with this algorithm passes through three stages, namely color space quantization (number reduction process of distinct colors in a given image), hit rate regions and similar color regions merging.

In the first stage, the color space is quantized with little perceptual degradation by using the quantization algorithm (Deng *et* al, 1999b) with minimum coloring. Each color is associated with a class. The original image pixels are replaced by classes to form the class maps in the next stage. Before performing the hit rate regions, the J-image - a class map for each windowed color region, whose positive and negative values represent the edges and textures of the processing image - must be created with pixel values used as a similarity algorithm for the hit rate region. These values are called „J-values" and are calculated from a window placed on the quantized image, where the J-value belongs.
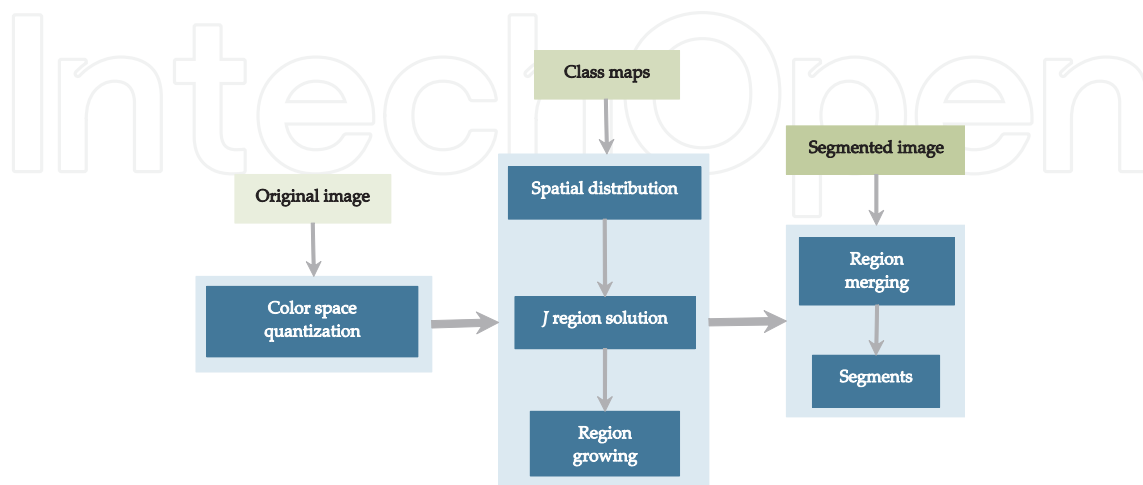
Figure 1. JSEG image segmentation steps.

2.1. Segmentation algorithm evaluation

Natural scenes present a 24-bit chromatic resolution color image, which is coarsely quantized preserving its major quality. The main idea for a good segmentation criterion is to extract representative colors differentiating neighboring regions in the acquired image, as an unsupervised method.

Therewith, the color quantization using peer group filtering (Deng *et al.*, 2001) is applied through perceptual weighting on individual pixels, to smooth the image and remove the existing noise. Then, new values indicating the smoothness of the local areas are obtained, and a weight is assigned to each pixel, prioritizing textured areas to smooth areas. These areas are identified with a quantization vector to the pixel colors, based on General Lloyd Algorithm (GLA) (Gersho and Gray, 1999), which the perceptually uniform L*u*v color space is adopted, presenting the overall distortion D:

$$D = \sum_i D_i = \sum_i \sum_n v(n) \|x(n) - c_i\|^2 \otimes x(n) \hat{I} \; C_i \tag{1}$$

And it is derived for:

$$c_i = \frac{\sum v(n)x(n)}{\sum v(n)} \otimes x(n) \hat{I} \; C_i \tag{2}$$

The parameters: c_i is the centroid of cluster C_i, $x(n)$ and $v(n)$ are the color vector and the perceptual weight for pixel n. D_i is the total distortion for C_i.

With the centroid value, as denoted by Equation (2) - after the vector quantization and merged clusters, pixels with the same color have two or more clusters, affected by GLA global distortion. For merging close clusters with minimum distance between preset thresholds for two centroids, an agglomerative clustering algorithm is performed on c_i (Duda and Hart, 1970), as the quantization parameter needed for spatial distribution.

After clustering merging for color quantization, a label is assigned for each quantized color, representing a color class for image pixels quantized to the same color. The image pixel colors are replaced by their corresponding color class labels, creating a class-map.

In order to calculate the J-value, Z is defined as the set of all points of quantized image, then $z = (x, y)$ with $z \in Z$ and being m the average in all Z elements. C is the number of classes obtained in the quantization. Then Z is classified into C classes, Z_i are the elements of Z belonging to class i, where $i=1,...,C$, and m_i are the element averages in Z_i.

$$m_i = \frac{1}{N_i} \sum_{z \hat{I} Z} z \tag{3}$$

The J-value is as follows:

$$J = \frac{S_B}{S_W} = \frac{(S_T - S_W)}{S_W} \tag{4}$$

where:

$$S_T = \sum_{z \in Z} \left\| z - m \right\|^2 \tag{5}$$

$$S_W = \sum_{i=1}^{C} \sum_{z \in Z} \left\| z - m_i \right\|^2 \tag{6}$$

The parameter S_T represents the sum of quantized image points within the average in all Z elements. Thereby, the relation between S_B and S_W, denotes the measures of distances of this class relation, for arbitrary nonlinear class distributions. J for higher values indicates an increasing distance between the classes and points for each other, considering images with homogeneous color regions. The distance and consequently, the J value, decrease for images with uniformly color classes.

Each segmented region could be recalculated, instead of the entire class-map, with new parameters adjustment for $\bar{J}$ average. J_K represents J calculated over region k, M_k is the number of points in region k, N is the total number of points in the class-map, with all regions in class-map summation.

$$\bar{J} = \frac{1}{N} \sum_{k} M_k J_k \tag{7}$$

For a fixed number of regions, a criterion for $\bar{J}$ is intended for lower values.

2.2. Spatial segmentation technique

The global minimization of $\bar{J}$ is not practical, if not applied to a local area of the class-map. Therefore, the idea of *J-image* is the generation of a gray-scale image whose pixel values are the J values calculated over local windows centered on these pixels. With a higher value for J-image, the pixel should be near region boundaries.

Expected local windows dimensions determines the size of image regions, for intensity and color edges in smaller sizes, and the opposite occurs detecting texture boundaries.

Using a region-growing method to segment the image, this one is considered initially as one single region. The algorithm for spatial segmentation starts segment all the regions in the image at an initial large scale until the minimum specified scale is reached. This final scale is settled manually for the appropriate image size. The initial scale 1 corresponds to 64x64 image size, scale 2 to 128x128 image size, scale 3 to 256x256 image size, with due proportion for increasing scales and the double image size.

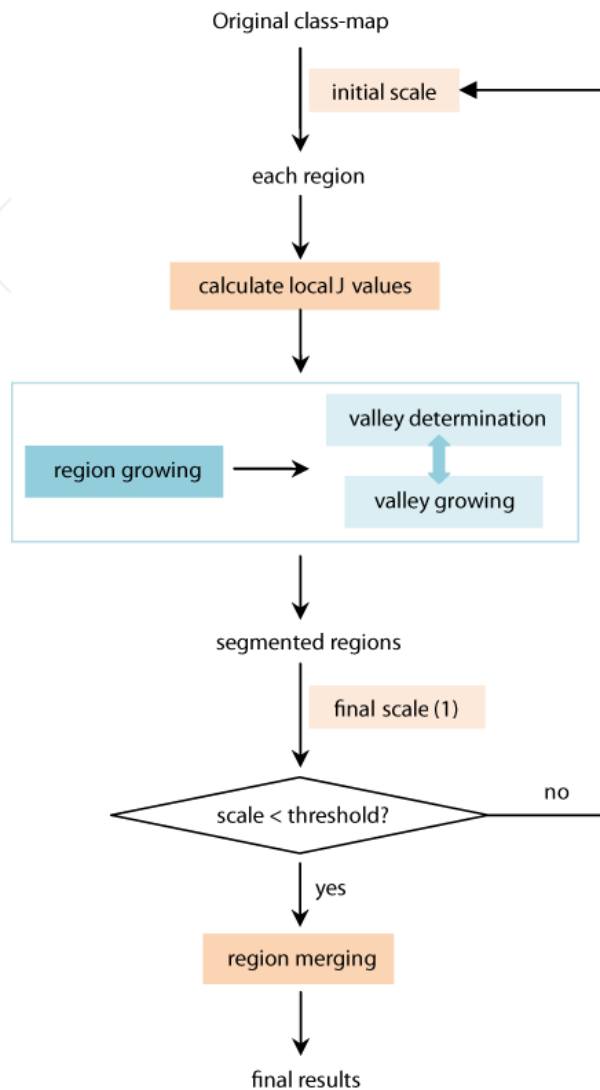Below, the spatial segmentation algorithm is structured in flow steps.

Figure 2. Sequence for spatial segmentation algorithm.

3. Image processing (spatial distribution and objects quantification)

The sequential images evince not only the color quantization (spatial distributions forming a map of classes), but also the space segmentation (J-image representing edges and regions of textured side).

Several window sizes are used by J-values: the largest detects the region boundaries by referring to texture parameters; the lowest detects changes in color and/or intensity of light. Each window size is associated with a scale image analysis. The concept of J-image, together with different scales, allows the segmentation of regions by referring to texture parameters.

Regions with the lowest values of J-image are called valleys. The lowest values are applied with a heuristic algorithm. Thus, it is possible to determine the starting point of efficient growth, which depends on the addition of similar valleys. The algorithm ends when there are spare pixels to be added to those regions.
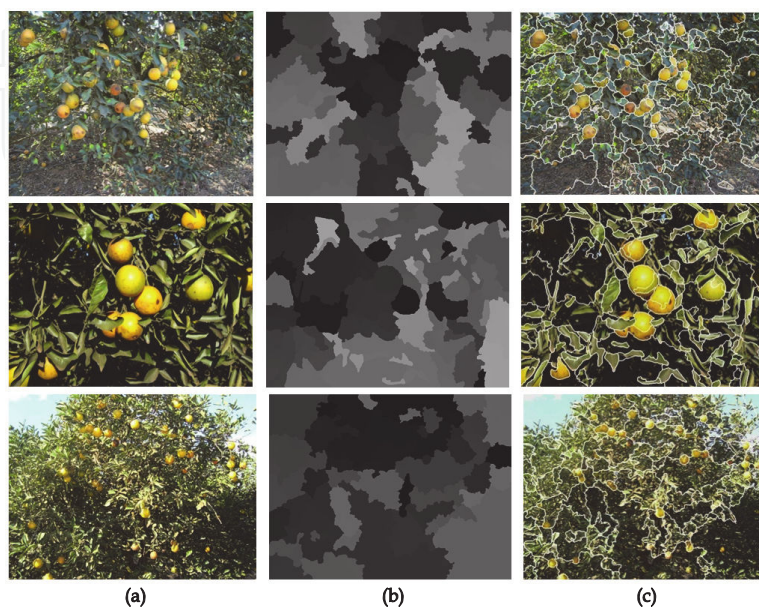
Figure 3. a) Original images; (b) Color quantization (map of classes); (c) J-image representing edges and regions of textured side (Spatial distributions).

It was observed that the oranges represent the largest number of image pixels, given its characteristics of high contrast with other objects on the scene.

Fig. 3, above, shows three types of scenes in orchards. The first identifies the largest part of the tree. In this category, the quantization threshold was adjusted to higher values for the fusion of regions with same color tone between branches, leaves and ground would be avoided. The second scene denotes the regions' set details in orchards, excluding darker regions. Not only irregularities of each leaf are segmented, as well as abnormalities of color tones in fruit itself, allowing later analysis of disease characteristics. The third category identifies most of the trees, but with higher incidence of top and bottom regions.

4. Artificial Neural Networks (ANN) – MLP customized algorithm

It is fundamental that an ANN-based classification method associated with a statistical pattern recognition be used. *Multi-Layer Perceptron* (MLP) (Haykin, 1999; Haykin, 2008) is suitable for default ANN topology to be implemented through a customized *back-propagation* algorithm for complex patterns (Costa and Cesar Junior, 2001).

The most appropriate segment and topology classifications are those using vectors extracted from HSV color space (Hue, Saturation, Value), matching RGB color space (Red, Green,

Blue) components. Also, the network with less MSE in the neurons to color space proportion is used to classify the entities.
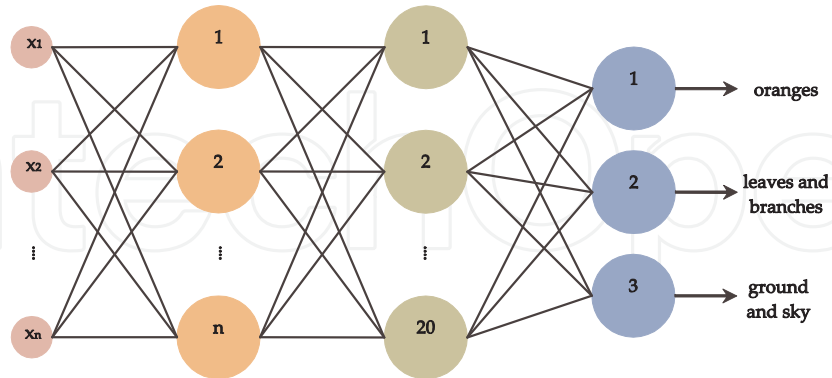

Figure 4. ANN schematic topology for fruits with three classes.

Derived from back-propagation, the iRPROP algorithm (improved resilient back-propagation) (Lulio, 2010) is both fast and accurate, with easy parameter adjustment. It features an Octave (Eaton, 2006) module which was adopted for the purposes of this work and it is classified with HSV (H – hue, S – saturation, V – value) color space channels histograms of 256 categories (32, 64,128 and 256 neurons in a hidden layer training for each color space channel: H, HS, and HSV). The output layer has three neurons, each of them having a predetermined class.
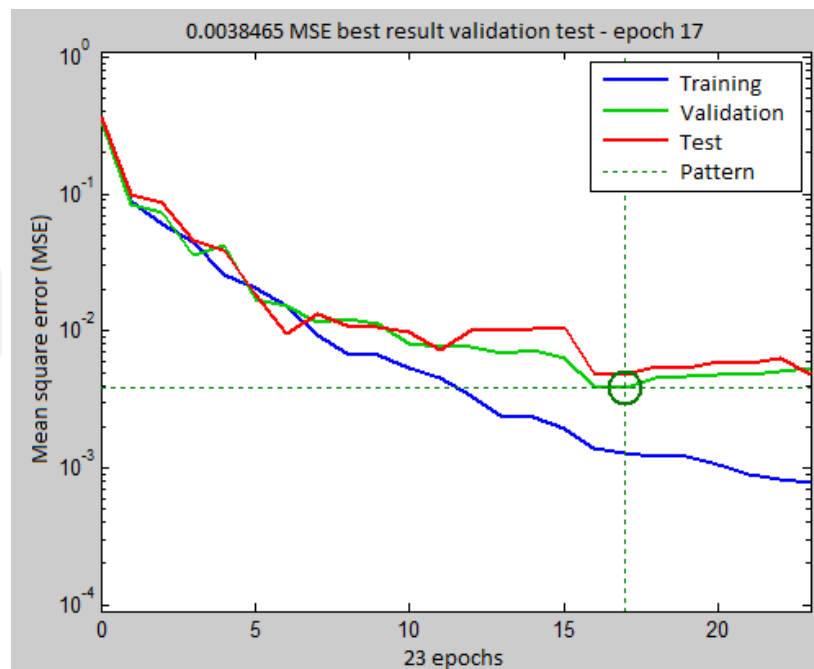
Figure 5. MSE 50% validation tests for RGB.

The charts below (Figures 5, 6, 7, 8) denote the ratio of mean square error (MSE) and amount of times to obtain the best performance index during the validation data towards the training and test sets.

All ANN-based topologies are trained with a threshold lower than 0.0001 mean squared errors (MSE), the synaptic neurons weights are initiated with random values and the other algorithm parameters were set with Fast Artificial Neural Network (FANN) library (Nissen, 2006) for Matlab (Mathworks Inc.) platform, and also its Neural Network toolbox. The most appropriate segment and topology classifications are those using vectors extracted from HSV color space. Also, a network with less MSE in the H-64 was used so as to classify the planting area; for class navigable area (soil), HSV-256 was chosen; as for the class sky, the HS-32.
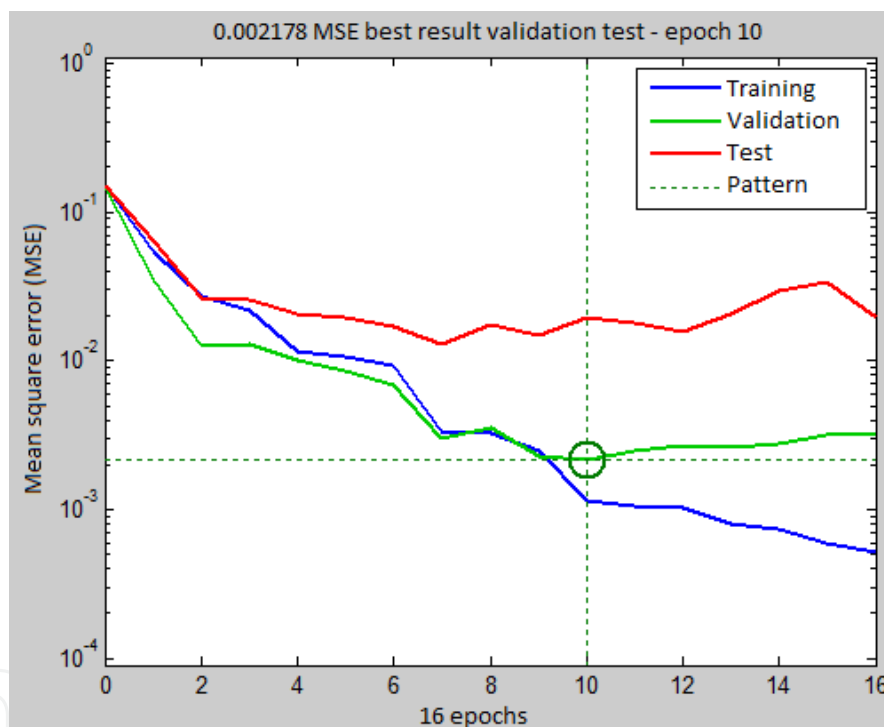
Figure 6. MSE 100% validation tests for RGB.

Figures 9 and 10 denote the regression for target-outputs of ANN classifier, for RGB and HSV classes. The higher the concentration of data at the intersection of bias and Y = T (equal to the output sampling period), the lower the linear regression of data is classified, based on confusion matrices for each set of dimensions.

The response times are given for combinations of training, testing, validation and all data sets.
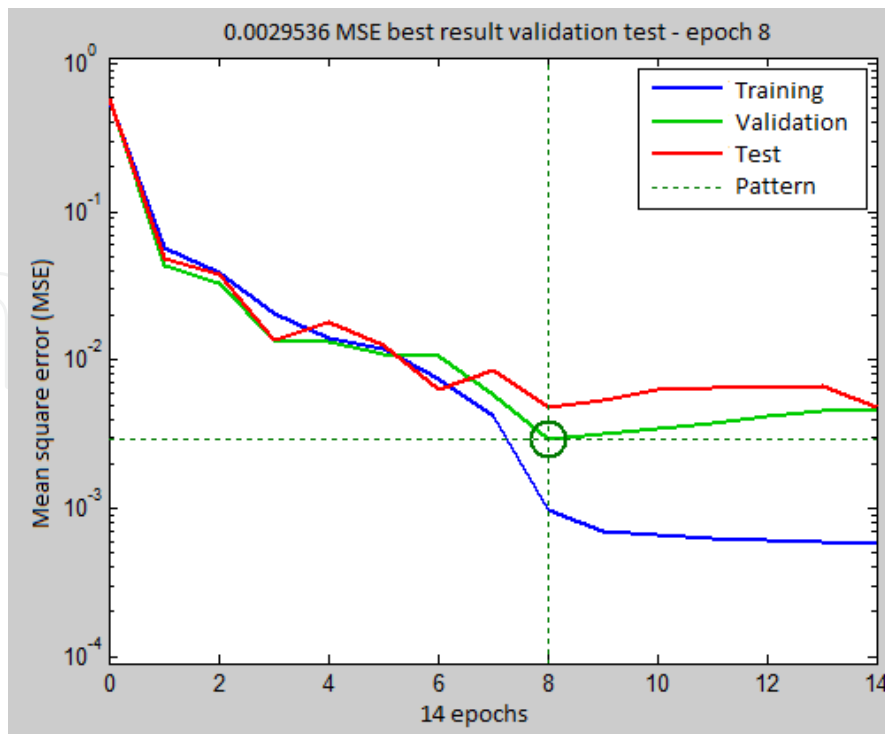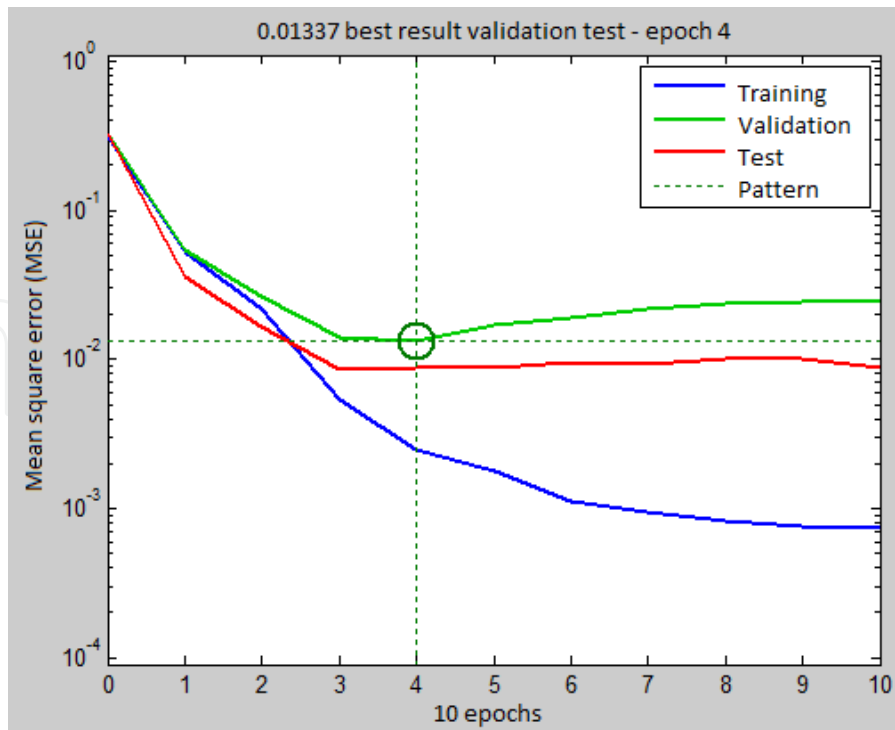
Figure 7. MSE 50% validation tests for HSV.

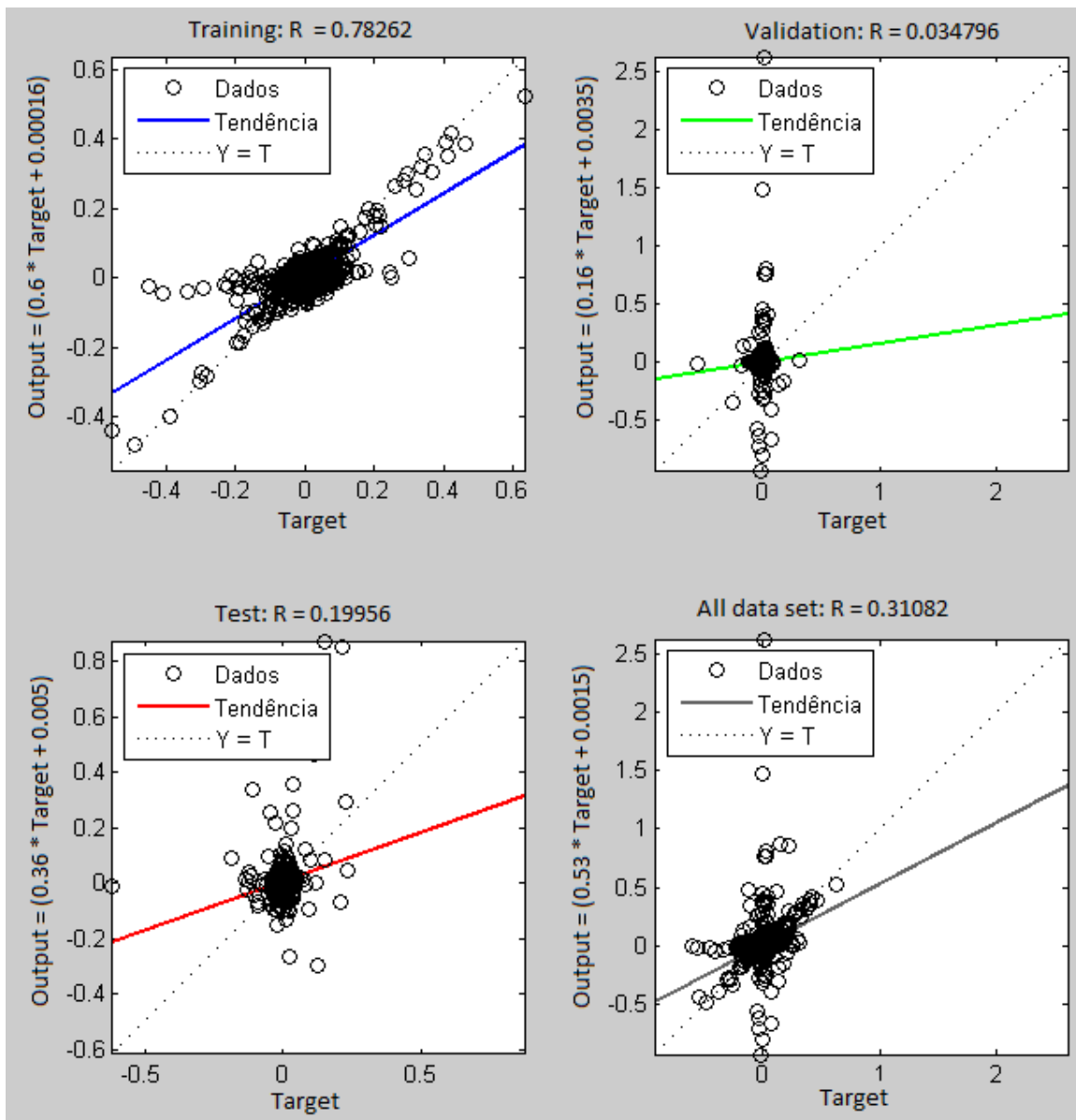**Figure 8.** MSE 100% validation tests for HSV.

Figure 9. MSE 50% (left) and 100% (right) validation tests for RGB.

5. Statistical pattern recognition

Statistical methods are employed as a combination of results with ANN, showing how accuracy in non-linear features vectors can be best applied in a MLP algorithm with a statistical improvement, which processing speed is essentially important, for pattern classification. *Bayes Theorem* and *Naive Bayes* (Comaniciu and Meer, 1997) both use a technique for iterations inspection, namely MCA (*Main Component Analysis*), which uses a linear transformation that minimizes co-variance while it maximizes variance. Features found through this transformation are totally uncorrelated, so the redundancy between them is avoided. Thus,

the components (features) represent the key information contained in data, reducing the number of dimensions. Therefore, RGB space color is used to compare the total number of dimensions in feature vectors with HSV. With a smaller dimension of iterations, HSV is chosen as the default space color in most applications (Grasso and Recce, 1996).
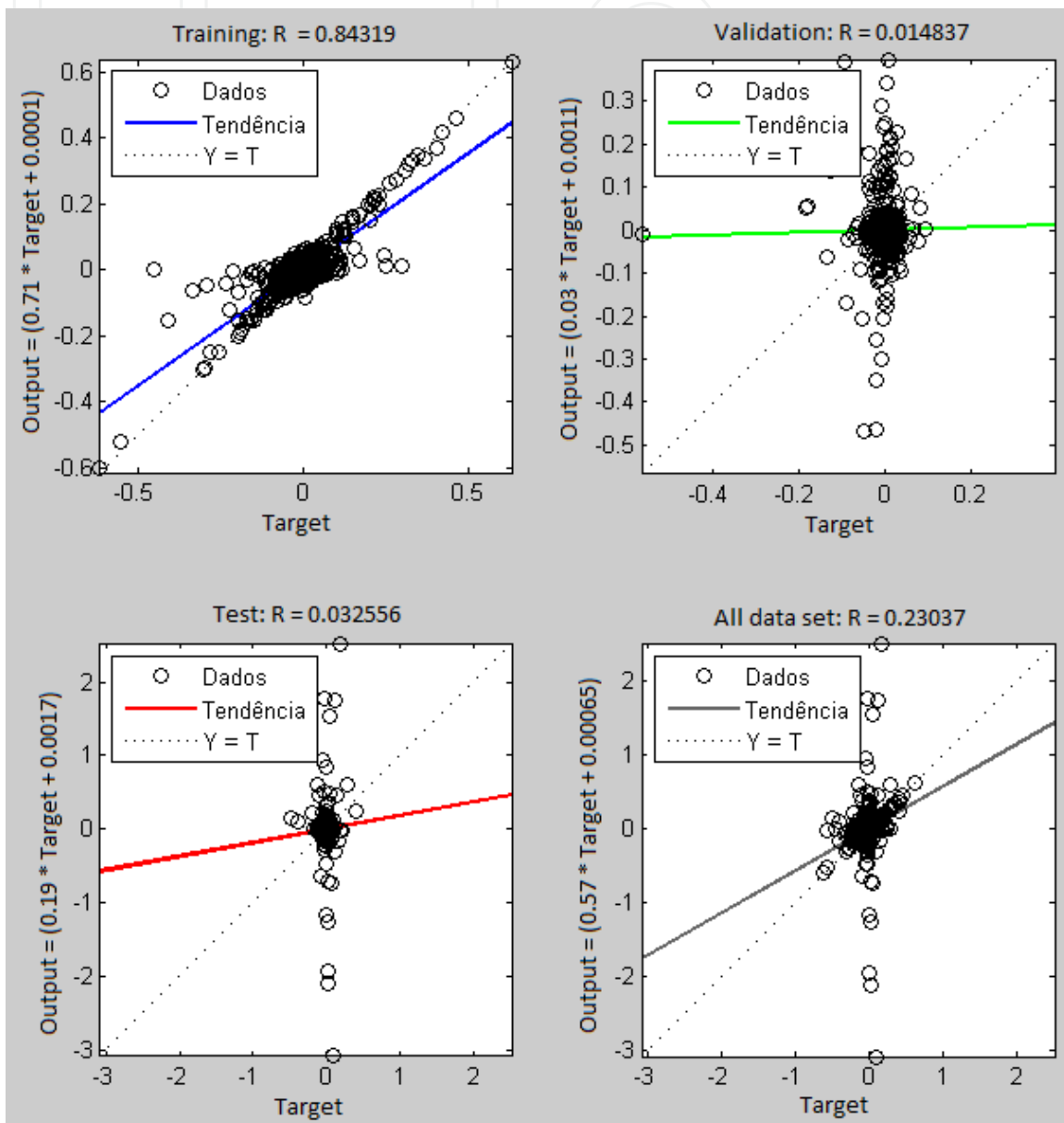

Figure 10. MSE 50% (left) and 100% (right) validation tests for RGB.

Bayes Theorem introduces a modified mathematical equation for the Probability Density Function (PDF), which estimates the training set in a conditional statistics. Equation (8) denotes the solution for $p(C_i|y)$ relating the PDF to conditional class i (classes in natural scene), and y is a n-dimensional feature vector. *Naive Bayes* implies independence for vector fea-

tures, what means that each class assumes the conditional parameter for the PDF, following Equation (9) (Morimoto *et* al, 2000).

$$P(C_i \mid y) = \frac{p(y \mid C_i)P(C_i)}{\sum_{j=1}^{K} p(y \mid C_j)P(C_j)} \tag{8}$$

$$P(y \mid C_i) = \prod_{j=1}^{u} p(y_j \mid C_i) \tag{9}$$

In Fig. 11, for the location of fruits in the RGB case, the discrimination of the classes fruit, sky and leaves, twigs and branches, attends constant amounts proportional to the increasing of the training sets. This amount, for HSV case, is reduced for the fruit class, as the dispersion of pixels is greater in this color space. In Fig. 12, in the RGB case, the best results were obtained using Bayes classifier, having smaller ratio estimation in relation to the number of components analyzed. In this color space, the estimation in the recognition of objects related to the fruits is given by the PDF of each dimension, correcting the current values by the hope of each area not matched to the respective class.

Also in Fig. 12, the recognition of the fruit to the HSV case presents balance in the results of the two classifiers, but with a compensation of the success rate, for lower margins of the estimation ratio to the Bayes classifier. This allows the correction of the next results by priori estimation approximating, in the PDF of each dimension.

It can be seen that, the ratio of the estimation must be lesser for the increasing of the dimensions number and its subsequent classification, in all cases.
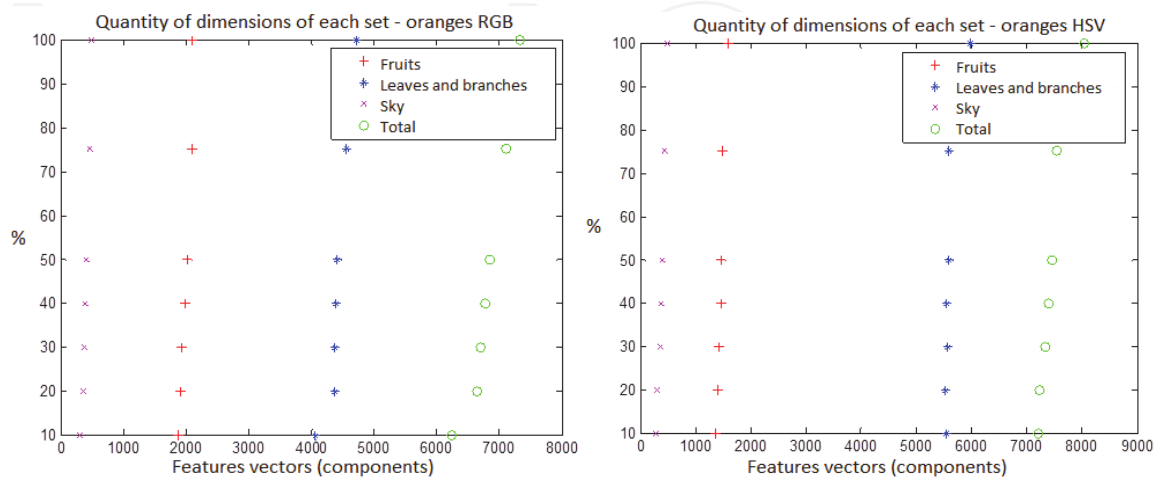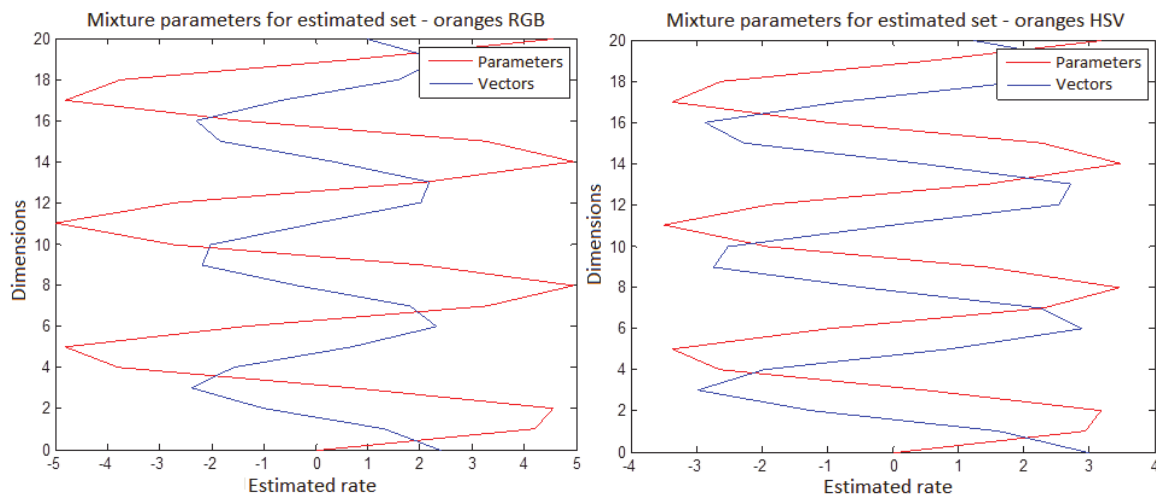

Figure 11. Quantity of dimensions of each set (oranges RGB - left, oranges HSV - right).

Figure 12. Mixture parameters for estimated set (oranges RGB - left, oranges HSV - right).

6. Objects quantification (post-processing)

The classes maps are processed, as the representation by the area filling (*floodfill*) brings only solid regions which are quantified. Initially, a conversion is performed on gray level image in order to threshold regions that are outlined. Then, to determine the labels of the elements connected, it is necessary to exclude objects which are greater than 200 to 300 pixels, depending on the focal length. Thus, it is necessary to identify each element smaller than this threshold, and calculate the properties of these objects, such as area, centroid, and the boundary region. As a result, the objects that present areas near the circular geometry will be labelled and quantified as fruits.

To determine the metrics and the definition of objects of orange crop, the graph-based segmentation (Gonzalez and Woods, 2007) was applied. This technique provides the adjacency relation between the binary values of the pixels, and their respective positions, highlighting the local geometric properties of the image.

In first case, areas corresponding to small regions, as fruits partially hidden (oranges) with equivalent texture and color properties to leaves are excluded. Then, estimated elements are fully grouped, when overlap the representative segments, which denote an orange fruit. Lastly, the grouping is applied for regions which detect two or more representative segments, denoting another orange fruit.

As the best classification results, related to second approach were through Bayes in HSV color space, only the maps of class from these classifiers will be presented to localization and quantification of objects, compared to RGB case.

Then, for the RGB and HSV cases are presented, through Figures 13 to 21, the images in their respective maps of class, the pre-processing for thresholding with areas smaller than
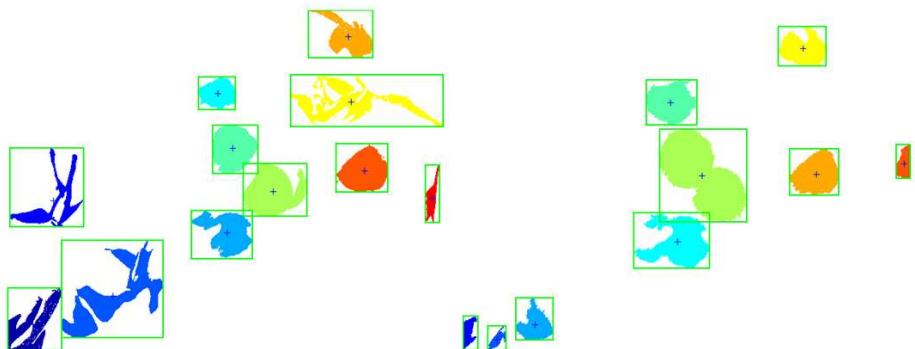
100 and greater than 300, the geometric approximation metrics for the detection of circular objects, the boundary regions with the centroid of each object, and finally the label associated to the fruit.

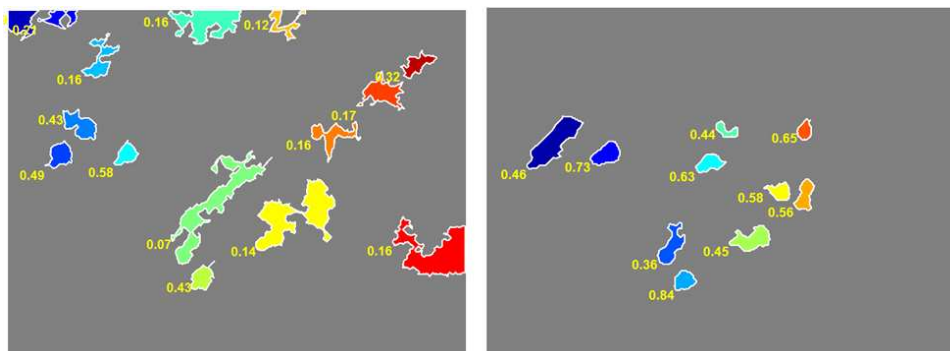**Figure 13.** Maps of RGB (left) and HSV (right) classes - scene 1.
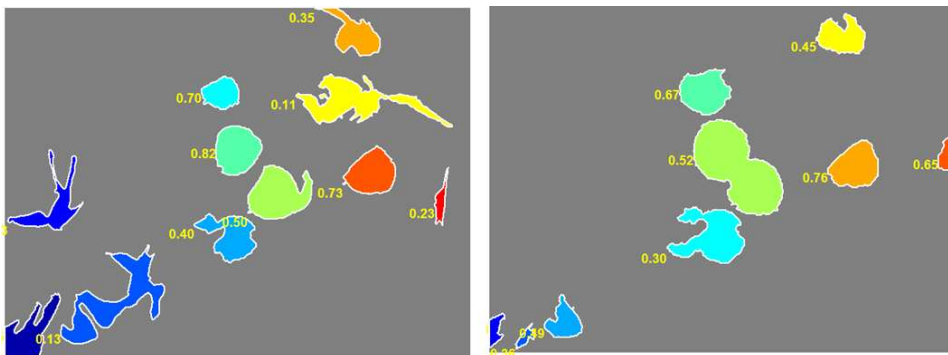
Figure 14. Metric near circular geometry threshold 1.0 for RGB (left) and HSV (right) - scene 1.
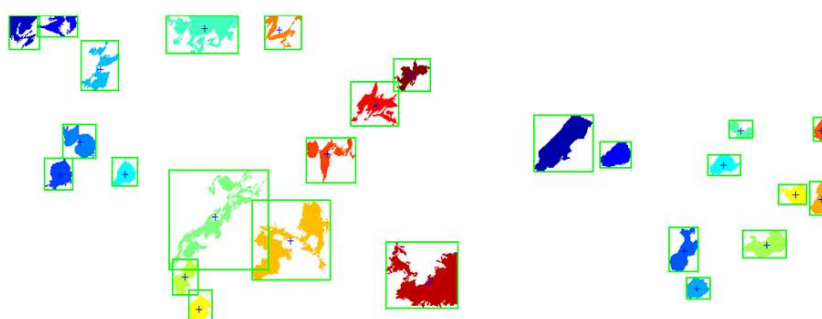

Figure 15. Representation of area and centroid for fruit association in two cases - scene 1.

Figure 16. Maps of RGB (left) and HSV (right) classes - scene 2.

Figure 17. Metric near circular geometry threshold 1.0 for RGB (left) and HSV (right) - scene 2.

Figure 18. Representation of area and centroid for fruit association in two cases - scene 2.

Figure 19. Maps of RGB (left) and HSV (right) classes - scene 3.

Figure 20. Metric near circular geometry threshold 1.0 for RGB (left) and HSV (right) - scene 3.

Figure 21. Representation of area and centroid for fruit association in two cases - scene 3.
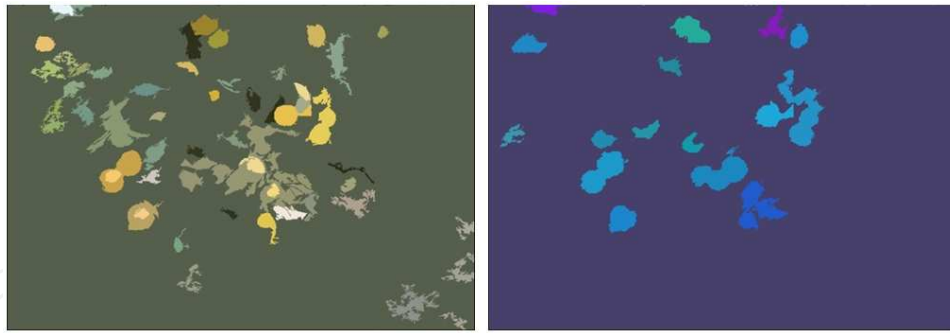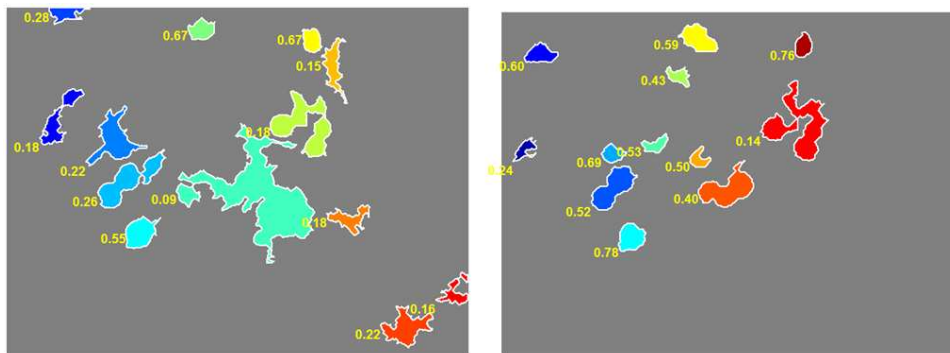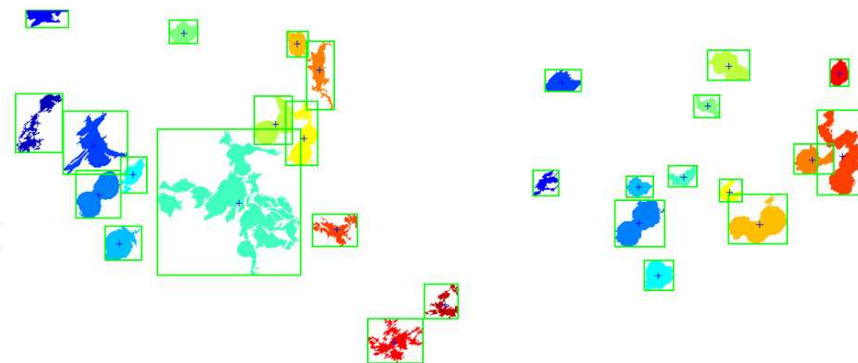
7. Conclusions

This chapter presented merging techniques for segmentation and statistical classification of agricultural orange crops scenes, running multiple segmentation tests with JSEG algorithm

possible. As the data provided evince, this generated algorithms fulfills the expectations as far as segmenting is concerned, so that it sorts the appropriate classes (fruits; leaves and branches; sky). As a result, a modular strategy with Bayes statistical theorem can be an option for the classification of segments applied with cognitive approach.

Author details

Luciano Cássio Lulio[1], Mário Luiz Tronco[1], Arthur José Vieira Porto[1],
Carlos Roberto Valêncio[2] and Rogéria Cristiane Gratão de Souza[2]

1 Engineering School of Sao Carlos, University of Sao Paulo (EESC/USP), Sao Carlos, São Paulo, Brazil

2 Statistical and Computing Science Department, State University of Sao Paulo (DCCE/ UNESP), São Jose do Rio Preto, São Paulo, Brazil

References

[1] Comaniciu, D., & Meer, P. (1997). Robust analysis of feature spaces: color image segmentation. In: Conference on Computer Vision and Pattern Recognition, IEEE Computer Society

[2] Costa, L. F., & Cesar, Junior. R. M. (2001). Shape analysis and classification- Theory and Practice. 1. ed. Boca Raton, Florida, EUA: CRC Press LLC. 0-84933-493-4

[3] Deng, Y., Kennedy, C., Moore, M. S., & Manjunath, B. S. (1999a). Peer group filtering and perceptual color image quantization. Proceedings of the 1999 IEEE International Symposium on Circuits and Systems, , 4, 21-25.

[4] Deng, Y., Manjunath, B. S., & Shin, H. (1999b). Color image segmentation. Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, , 2, 446-451.

[5] Deng, Y., & Manjunath, B. S. (2001). Unsupervised segmentation of color-texture regions in images and videos. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI'01), , 23(8), 800-810.

[6] Duda, R. O., & Hart, P. E. (1970). Pattern Classification and Scene Analysis, John Wiley & Sons, New York

[7] Eaton, J. W., et al. (2006). Octave. Avaliable at: http://www.octave.org

[8] Gersho, A., & Gray, R. M. (1992). Vector quantization and signal compression, Kluwer Academic, Norwell, MA

[9]   Gonzalez, R. C., & Woods, R. E. (2007). Digital Image Processing. 3 ed. New Jersey, EUA: Prentice-Hall Inc

[10] Grasso, G. M., & Recce, M. (1996). Scene Analysis for an Orange Picking Robot,"In: International Congress for Computer Technology in Agri-culture

[11] Haykin, S. (1999). Neural networks: a comprehensive foundation. 2. ed. New Jersey, EUA: Prentice-Hall. 0-13273-350-1

[12] Haykin, S. (2008). Neural Networks and Learning Machines. 3.ed. McMaster University, Canada: Prentice-Hall. 0-13147-139-2

[13] Igel, C., & Hüsken, M. (2003). Empirical evaluation of the improved Rprop learning algorithm. Neurocomputing, , 50, 105-123.

[14] Lulio, L. C. (2011). Computer vision techniques applied to natural scenes recognition and autonomous locomotion of agricultural mobile robots. São Carlos, 353 p. Dissertation (Master of Science)- School of Engineering of São Carlos, University of São Paulo, São Carlos

[15] Morimoto, T., Takeuchi, T., Miyata, H., & Hashimoto, Y. (2000). Pattern recognition of fruit shapes based on the concept of chaos and neural networks,"Computers and Electronics in Agriculture, , 26, 171-186.

[16] Nissen, S., et al. (2006). Fann: fast artificial neural network library. Avaliable at: http://leenissen.dk/fann/