JSEG Algorithm and Statistical Image Segmentation Techniques for Quantization of Fruits

Luciano Cassio Lulio, Valdinei Luis Belini
Department of Mechanical Engineering
University of São Paulo, USP
São Carlos, Brazil
Email: mailto:luciano.lulio@gmail.com

Mario Luiz Tronco, Arthur Jose Vieira Porto
Department of Mechanical Engineering
University of São Paulo, USP
São Carlos, Brazil

***Abstract*—Computer vision, when used in open and unstructured environments as in the inspection of crops, requires the use of specific vision systems acquiring algorithms prepared for such situations. These algorithms work mainly with images composed of complex objects, textures, shadows and brightness. The present project aims to apply merging techniques as unsupervised homogeneous segmentation algorithm and statistical classification for mobile robot inspection in agricultural orange crops for quantization of fruits.**

***Keywords*-Segmentation; Fruit; Harvest;**

I. Introduction

Fruit recognition from images taken in an unstructured environment for fruit-harvesting robotic applications is a challenging task mainly due to the variable illumination, partial occlusion of the fruit area, shadow effect due to sunlight incident angle, merging of backgrounds (foliage), uncertainties due to shape of the fruits and the real-time restriction, just to mention some issues. According to [1], citrus are typical trees with a dense foliage and quite often 70-80% of the fruits have half or more of their surface obscured.

Despite these difficulties, several approaches based on image processing algorithms have been proposed to segment fruit from natural images [2], [3], [4], [5], [1], [6], and [7].

Other methods have also been proposed, such as those based on laser range-finder [8], the concept of chaos and neural networks [9], Support Vector Machine [10], and machine vision and ultrasonic sensors integration [11].

Comparative experiments involving Dynamic threshold segmentation method, extended Otsu, improved Otsu combined with genetic arithmetic, and adaptive segmentation method based on Learning Vector Quantization network was conducted by [12]. Tests using images of apple, tomato, strawberry, persimmon and orange demonstrated that Dynamic Threshold method has better performance and least cost time than extended Otsu method, improved Otsu combined with genetic arithmetic and adaptive segmentation method based on LVQ network. Besides, it has satisfactory effect upon fruit image under natural illumination condition. Adaptive segmentation method based on LQV network can only be applied into balanced color instance of particular fruit, and it is not adapt to be applied into real-time occasion because of high cost time [12].

Regarding fruit-harvesting image segmentation two main problems must be overcome. Firstly, the influence imposed by lighting condition, which depends on the incident solar angle on the scene. This causes a significant difference in how the harvesting scene appears. Furthermore, fruits inside the canopy receive a different amount of illumination compared with the fruits on the canopy surface. Thus, the image processing algorithm should be robust enough to deal with this kind of lighting variation. The second problem is related with occlusion, which minimizes the fruit area visibility and disrupts the shape of the fruit. This greatly affects the ability to detect oranges simply by their shape or size. The main causes of occlusion are leaves, branches and other fruits. Unlike leaf or branch occlusion where there is a sharp contrast in color between the fruit and leaf, fruit occlusion can cause multiple fruits to appear as a single fruit [13].

While the influence of the illumination can be alleviated by using, for instance, external (artificial) light sources and light compensation mechanisms, the occlusion phenomenon represents a more relevant problem once it can result in the under estimation of the amount of fruits in automatic counting algorithms.

Contributions: Motivated by this observation, the present work has deal with occlusion in the fruit natural image segmentation process by using a gradient-based method. This method provides the segmentation of homogeneous regions presented mainly in natural scenes. The results obtained by applying the Bayesian classifier demonstrated heterogeneously the identification of fruits in the scenes of orange orchards. The HSV color space shown to be superior to RGB given the natural normalization of the hue component in relation to the constituent components of the RGB color space.

The outline of the paper is organized as follows. Section II describes the used methodology. Experimental results are presented in Section III and conclusions are given in Section IV.

A. Related work

Many papers have already been proposed in order to reduce the influence of lighting conditions and occlusion on segmenting fruits from natural images. A quite complex sequence of processing steps was presented by [1] for guiding an orange picking robot. By projecting all of the pixels along a direction perpendicular to the normal HS plane (from the conventional HSV color space) (please, see [1] for details) a new color space was constructed to segment fruit images. This operation pushes pixels with a low saturation or low intensity close to the space center and pixels, which are more significant, possessing a large product of saturation and intensity far from the center. As a consequence, two well-separated clouds of pixels can be distinguished, one in the region of red-orange, which corresponds to the distribution of oranges color and the other crossing the origin horizontally from blue to green, which corresponds to the background. Thus, the overlap between the region of interest and the background in the modified HSV space is significantly reduced [4].

The classification of regions in the color space and segmentation of the image of an orange tree was performed using a back-propagation multilayer perceptron. To deal with fruits occurred in clusters, an adaptive edge-fitting algorithm has been adopted to extract orange centers. This algorithm calculates the area of overlap of each pair of identified circular regions and if this overlap is greater than a previously defined threshold the corresponding circle is rejected. The performance of the algorithm for orange detection from over 50 images containing 673 oranges was indicated by 87% of oranges correctly identified and 15% for false positive (leaf, for example, was considered an orange) and 5% for false negative(real orange was not included). Despite these promising results, the authors recognized that the proposed algorithm had significant problem with some types of occlusions, in which leaves were more of a problem than clusters of oranges. From 150 scenes, in only 13% of these scenes the success rate was about 90%.

An algorithm capable of counting number of oranges was developed by [6]. Image acquisition, binarization of color image in hue-saturation color plane, preprocessing to remove noise and to fill gaps. Due to the fact that fruit portion, leaf portion, and the background are not easily differentiated using traditional gray level threshold of global threshold [6] the authors has classified pixels in the HIS (Hue, Saturation, and Intensity) color space into three classes: citrus fruits, leaf, and background. These classes were chosen manually to the calibration process. By comparing RGB and HSI the authors found a clear line of separation between the fruits and the background in the second color space. After including the luminance component into a threshold carefully chosen in the calibration process the following pixel distribution for 25 images in HIS color plane was found: 99.7% and 100% for the lead class and background classes, respectively, were outside threshold. Although, only 58% do citrus class was captured inside the threshold, the succeeding binarization scheme performed satisfactory with validation images. The noises contained in the binarized images, mainly due to the little overlap of the lead class, were removed by adopting a threshold of 100 pixels.

As an attempt to avoid that single fruits occluded by small leaves were counted as more than one fruit, a set of dilation and erosion with a kernel size of 5x5 pixels was applied to the binarized images. In this context, citrus fruits were identified using blob analysis, in which fruit pixels were treated as a single blob and the total number of blobs gave the number of fruits in the image. In this method, any cluster of fruits was identified partially using the average area of a fruit and counted as two fruits. As a consequence, there were many under estimation. The authors argued that it has been occurred when some fruit clusters were counted as single fruit and when the fruit was removed by the threshold in area, designed to remove noise, erroneously interpreted small visible portion of a fruit as a noise. When comparing manual and automatic fruits counting on 329 validation images, a regression analysis indicated a value of 0.79.

The work in [7] presented a vision based algorithm capable of recognizing both on-tree red and green apples in a single RGB image frame. To achieve this goal they have combined texture property contrast from co-occurrence matrix and the color property redness. After calculating texture measure a Canny-based edge detection to contour extraction, whereas dilation and erosion was carried out to eliminate the effects of partial occlusion. The combination of a redness threshold into contoured regions was used to quality these areas to those corresponding to apples. To avoid that the algorithm confuse texture backgrounds with apple texture and redness, a Laplacian filtering was employed to provide edge enhancement. Thus, while this strategy enhanced contrast in texture (edge) to separate the background regions from apple regions the texture within the apples remained relatively unchanged. An extra benefit of using edge enhancing was shown in detecting apples at different distances from the image acquisition system. To separate clustered apples a circle fitting was employed.

A five image processing steps algorithm for orange detection was developed by [13]: (i) the segmentation stage explored a global thresholding based on red (normalization of the red component to the sum of the RGB channels) coordinates of the chromaticity diagram. The threshold value was determined manually; (ii) a labeling algorithm searches through the image looking for pixels that are connected to each other and, occasionally, define region corresponding to physical oranges. By using an important property defined by the fruit area, manually chosen, the (iii) size filtering step removes any region that is either too small or too large to be an orange. To locate orange fruits in the clusters determined by the labeling step a (iv) perimeter extraction process allows detecting (v) circles and their corresponding center candidates in each labeled region. The performance of the proposed method regarding the varying illumination and the occlusion was demonstrated with an average of 90% of the fruits correctly detected. The database for the tests was composed by images taken inside the canopy (with and without flash), near-view images (less than 0.5m) and far-view (images taken at 1.0 m distance).

II. Unsupervised homogeneous regions 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 electromagnetic radiation at many discrete wavelengths) [14]. 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.

An unsupervised color-texture regions segmentation algorithm is ideal for this purpose, since it tests the homogeneity of a given color-texture pattern, which is computationally more feasible than model parameter estimation. It deals with the following assumptions for the acquired image:

- Image containing homogeneous color-texture regions;
- Color information is represented by quantized colors;
- Colors between two neighboring regions are distinguishable.

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 two major stages, namely color space quantization (number reduction process of distinct colors in a given image), and hit rate regions with similar color regions merging, as secondary stage.

In the first stage, the color space is quantized with little perceptual degradation by using a quantization algorithm [15] [16] with minimum coloring. In this context, each color is associated with a class and the original image pixels are replaced by classes to form the class maps (texture composition) for 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. Therefore, the two-stage division is justified through the difficult analysis of the colors similarity withal their distributions.

The decoupling of these features (color similarity and spatial distribution) allows tractable algorithms development for each of the two processing stages (Fig. 1).

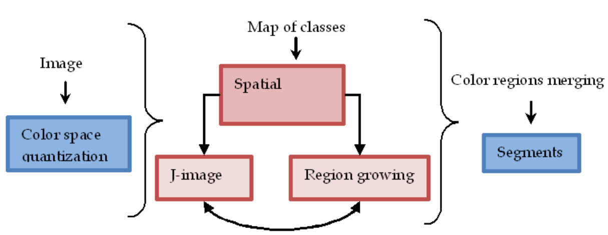

Fig. 1. JSEG image segmentation steps

A. 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 [17] 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) [18], which the perceptually uniform L*u*v color space is adopted, presenting the overall distortion D:

$$D = \sum_i D_i = \sum_i \sum_n \nu(n) \left\| x(n) - c_i \right\|^2 \longrightarrow x(n) \in C_i. \tag{1}$$

And it is derived for:

$$c_i = \frac{\sum \nu(n) x(n)}{\sum \nu(n)} \longrightarrow x(n) \in C_i. \tag{2}$$

The parameters: c_i is the centroid of cluster C_i, $x(n)$ and $\nu(n)$ 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 (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 [19], 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 Fig. 2, class-map 1 indicates three regions containing a single class of data points for segmentation process, and class-map 2 is not segmented indicating a color uniformity.

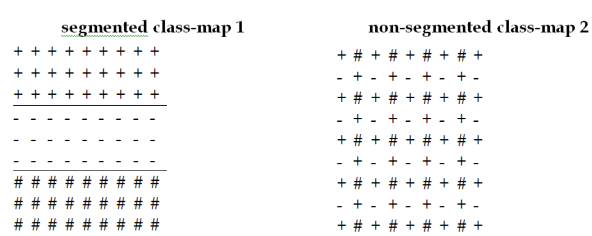

Fig. 2. Two different class-map representing three distinct classes of data points

The symbols ($+$, $-$, $\neq$) denote the label values (*J-value*) for three distinct data points. All necessary segmentation information, after color quantization, is extracted and relocated to a class-map. A specific region contains pixels from a color class set, which is distributed in image regions. These regions, forming each one, a class–map, have distributed points in all spatial data segments, corresponding a two-dimensional plane, and represent the cartesian position vector $(x,\ y)$.

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 average elements in Z_i.

$$m = \frac{1}{N}\sum_{z\in Z} z \tag{3}$$

$$m_i = \frac{1}{N_i}\sum_{z\in Z} z \tag{4}$$

The *J-value* is as follows:

$$J = \frac{S_B}{S_W} = \frac{(S_T - S_W)}{S_W} \tag{5}$$

$$S_T = \sum_{z\in Z} \|z - m\|^2 \tag{6}$$

where:

$$S_W = \sum_{i=1}^{C}\sum_{z\in Z} \|z - m_i\|^2 \tag{7}$$

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 $\overline{J}$ 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.

$$\overline{J} = \frac{1}{N}\sum_{k} M_k J_k \tag{8}$$

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

B. Segmentation Algorithm Evaluation

Bayes Theorem introduces a modified mathematical equation for the Probability Density Function (PDF), which estimates the training set in a conditional statistics. Equation (4) 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 features, what means that each class assumes the conditional parameter for the PDF, following Equation (9).

$$P(C_i|y) = \frac{p(y|C_i)P(C_i)}{\sum_{j=1}^{K} p(y|C_j)P(C_j)} \tag{9}$$

$$P(y|C_i) = \prod_{j=1}^{n} p(y_j|C_i) \tag{10}$$

III. Results and Discussion

A. Segmentation of the scenes of fruit orchards rows

Figure 3 shows three categories of orchards scenes. The first one identifies the most part of the tree, also evidencing the soil. In this category, the quantization threshold was adjusted to higher values in order to avoid the fusion of regions with the same color tonalities, e.g. stems, branches, twigs and soil. The second category represents details in given sets in the orchards scenes, excluding darker regions from the remaining scene. Besides the irregularities of each leaf the abnormalities of tones of the fruit are segmented, allowing posteriori analysis of characteristic diseases in oranges. The third category identifies, like the first, most of the trees, but with a higher incidence of crown and regions of the sky.

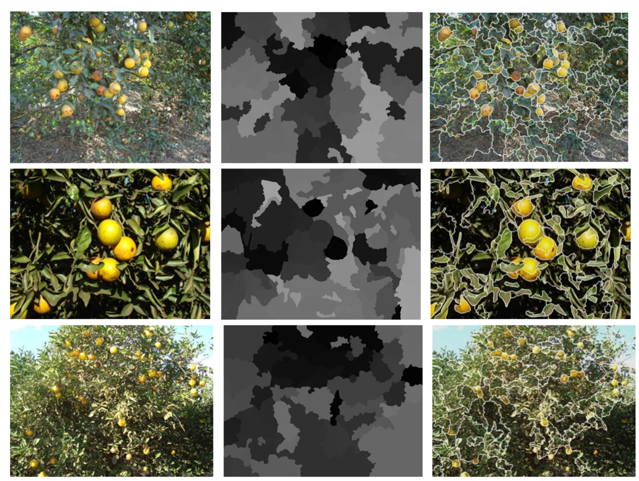

Fig. 3. Fruit orchards original scenes (column 1); quantized color images (column 2); segmented images (column 3)

In Fig. 4 and Fig. 5, 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.

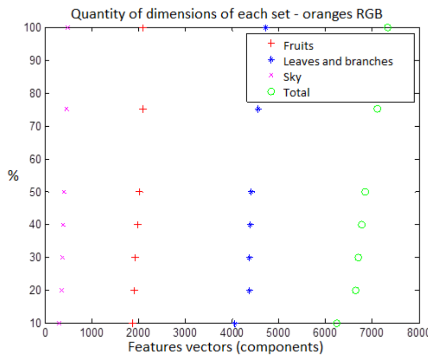

Fig. 4. Quantity of dimensions of each set (orange RGB)

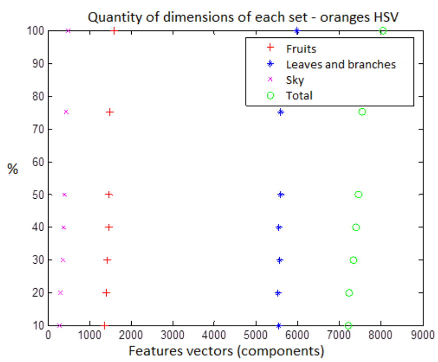

Fig. 5. Quantity of dimensions of each set (orange HSV)

In Fig. 6, for the second approach, 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.

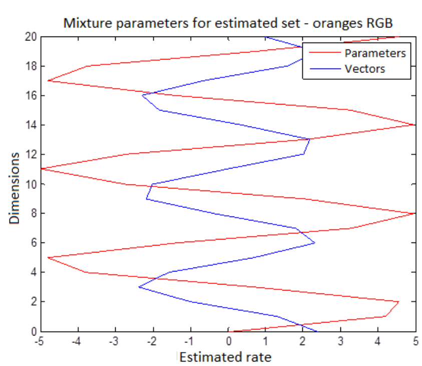

Fig. 6. Mixture parameters for estimated set (oranges RGB)

In Fig. 7, 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.

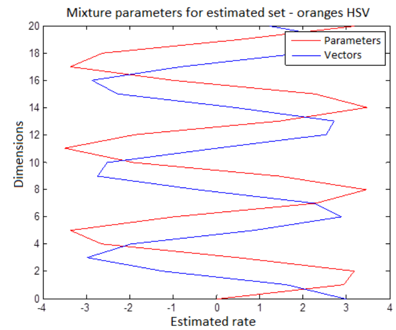

Fig. 7. Mixture parameters for estimated set (oranges HSV

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.

B. Post-processing in the objects quantification

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 with 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 labeled and quantified as fruits.

To determine the metrics and the definition of objects of orange crop, the graph-based segmentation [20] 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 Fig. 8 to 19, 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.

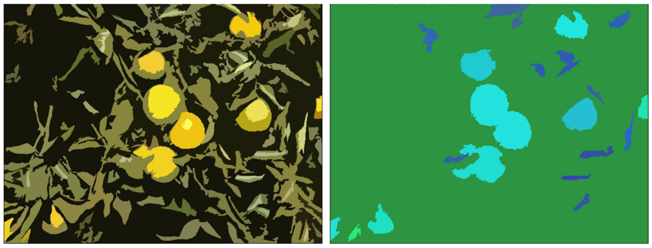

Fig. 8. Maps of RGB classe (left) and HSV (right) scene 1.

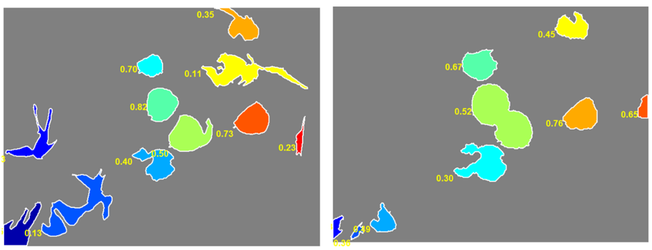

Fig. 9. Metric near the threshold circular geometric 1.0 for RGB (left) and HSV (right) — scene 1.

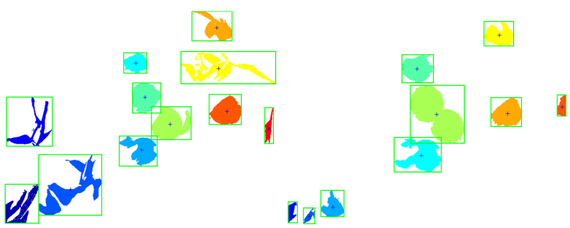

Fig. 10. Representation of the area and centroid of the objects to RGB (left) and HSV (right) — scene 1.

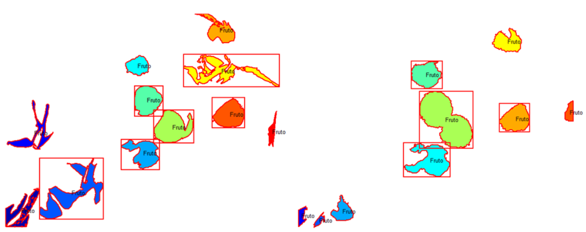

Fig. 11. Label of fruit element association in the found objects in two cases — scene 1.

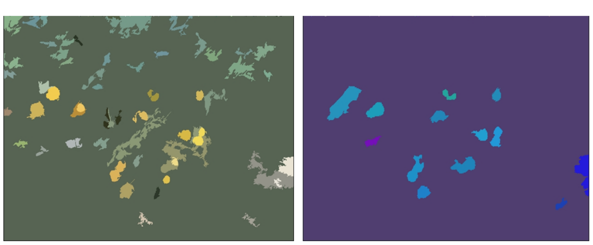

Fig. 12. Maps of class RGB (left) and HSV (right) — scene 2.

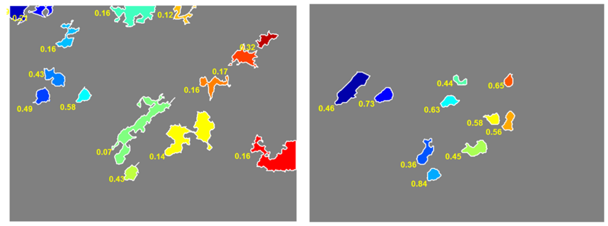

Fig. 13. Metrics near the threshold circular geometric 1.0 for RGB (left) and HSV (right) — scene 2.

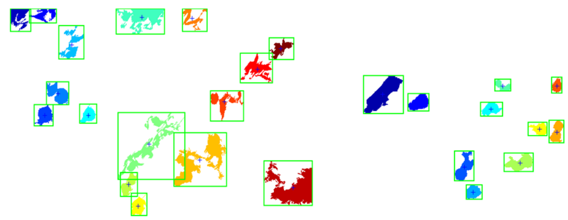

Fig. 14. Representation of the area and centroid of the objects to RGB (left) and HSV (right) — scene 2.

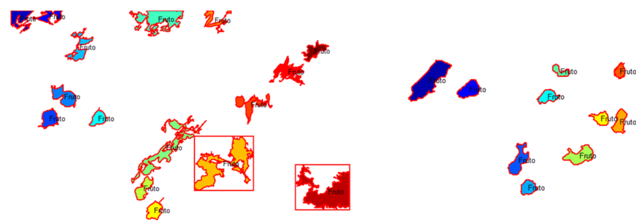

Fig. 15. Label of fruit element association in the found objects in two cases — scene 2.

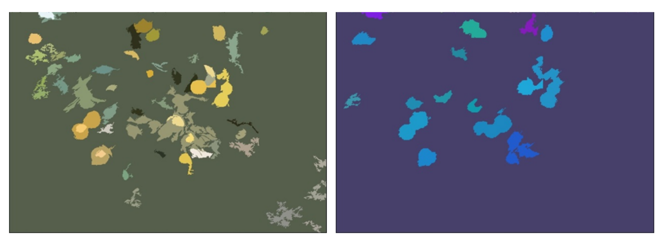

Fig. 16. Maps of class RGB (left) and HSV (right) — scene 3.

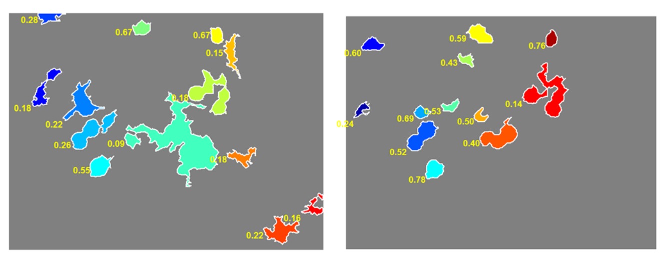

Fig. 17. Metrics near circular geometric threshold 1.0 for RGB (left) and HSV (right) — scene 3.

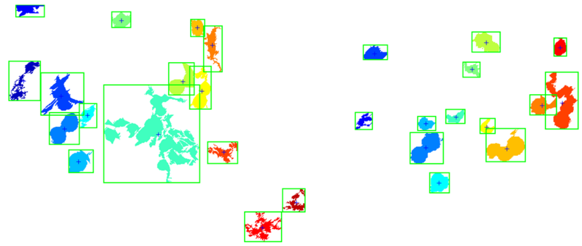

Fig. 18. Representation of the area and centroid of the objects to RGB (left) and HSV (right)−scene 3.

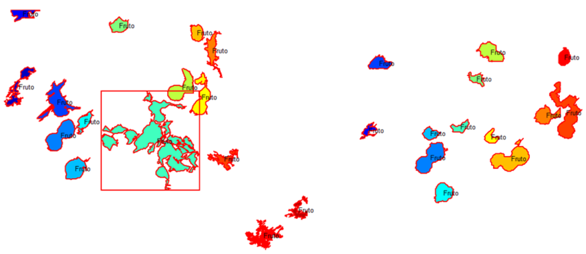

Fig. 19. Label of fruit element association in the found objects in two cases − scene 3.

IV. Conclusion

This paper 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. Moreover, the classification using different types of feature vectors caused the classification metric to be more accurate and sophisticated with Bayes, as well as the HSV color space to have lower MSE in test values for the post-processing stage.

References

[1] A. Plebe and G. Grasso, "Localization of spherical fruits for robotic harvesting," *Machine Vision and Applications*, vol. 13, pp. 70–79, 2001.

[2] D. C. Slaughter and R. C. Harrel, "Discriminating fruit for robotic harvest using color in natural outdoor scenes," *Transactions of the ASAE*, vol. 32, no. 2, pp. 757–763, 1989.

[3] F. Pla, F. Juste, and F. Ferri, "Feature extraction of spherical objects in image analysis: an application to robotic citrus harvesting," *Computers and Electronics in Agriculture*, vol. 8, no. 1, pp. 57 – 72, 1993.

[4] G. M. Grasso and M. Recce, "Scene Analysis for an Orange Picking Robot," *In: International Congress for Computer Technology in Agriculture*, 1996.

[5] A. R. Weeks, A. Gallagher, and J. Eriksson, "Detection of Oranges from a ColorImage of an Orange Tree," *In: Proceedings of SPIE - the International Society for Optical Engineering*, pp. 346–357, 1999.

[6] P. Annamalai, W. S. Lee, and T. F. Burks, "Color vision system for estimating citrus yield in real-time," *In: ASAE/CSAE Annual International Meeting*, 2004.

[7] J. Zhao, J. Tow, and J. Katupitiya, "On-tree fruit recognition using texture properties and color data," *IEEE/RSJ Int. Conf. Intell. Robots and Systems*, pp. 3993–3998, 2005.

[8] A. R. Jimenez, R. Ceres, and J. L. Pons, "A vision system based on a laser range-finder applied to robotic fruit harvesting," *Machine Vision and Applications*, pp. 321–329, 2000.

[9] T. Morimoto, T. Takeuchi, H. Miyata, and Y. Hashimoto, "Pattern recognition of fruit shapes based on the concept of chaos and neural networks," *Computers and Electronics in Agriculture*, vol. 26, pp. 171–186, 2000.

[10] J. J. Wang, J. W. Zhao, De-an, J. J. Tu, and Y. Zhang, "Application of support vector machine to apple recognition using in apple harvesting robot," *Proceedings of the 2009 IEEE International Conference on Information and Automation*, pp. 1110–1115, 2009.

[11] M. Regunathan and W. S. Lee, "Citrus Fruit Identification and Size Determination Using Machine Vision and Ultrasonic Sensors," *In: American Society of Agrigultural Engineers*, 2005.

[12] J. Yin, M. Hanping, and X. Yongliang, "Segmentation methods of fruit image and comparative experiments," *International Conference on Computer Science and Software Engineering. In: IEEE Computer Society*, pp. 1098–1102, 2008.

[13] M. W. Hannan, T. F. Burks, and D. M. Bulanon, "A Real-time machine vision algorithm for robotic citrus harvesting," *American Society fo Agricultural and Biological Engineers. In: An ASABE Meeting Presentation*, 2007.

[14] L. Lulio, "Computer vision techniques applied to natural scenes recognition and autonomous locomotion of agricultural mobile robots, Master thesis," Brazil, 2011.

[15] Y. Deng, B. Manjunath, and H. Shin, "Color image segmentation," *Conference on Computer Vision and Pattern Recognition, IEEE Computer Society*, vol. 2, pp. 446–451, 1999b.

[16] Y. Deng and B. Manjunath, "Unsupervised segmentation of color-texture regions in images and videos," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI '01)*, vol. 23, no. 8, pp. 800–810, 2001.

[17] Y. Deng, C. Kennedy, M. Moore, and B. Manjunath, "Peer group filtering and perceptual color image quantization," *Proceedings of the 1999 IEEE International Symposium on Circuits and Systems*, vol. 4, pp. 21–25, 1999a.

[18] A. Gersho and R. Gray, *Vector quantization and signal compression*. Kluwer Academic, Norwell, MA, 1992.

[19] R. Duda and P. Hart, *Pattern Classification and Scene Analysis*. John Wiley-Sons, New York, 1970.

[20] R. C. Gonzalez and R. E. Woods, *Digital Image Processing. 3 ed*. New Jersey, EUA: Prentice-Hall Inc, 2007.