COMPUTER VISION TECHNIQUES APPLIED TO NATURAL SCENES RECOGNITION AND AUTONOMOUS LOCOMOTION OF AGRICULTURAL MOBILE ROBOTS

Luciano Cassio Lugli, Mario Luiz Tronco, Arthur José Vieira Porto

lugli@usp.br; mltronco@sc.usp.br; ajvporto@sc.usp.br

Department of Mechanical Engineering, São Carlos School of Engineering

University of São Paulo

A paper from the Proceedings of the

14th International Conference on Precision Agriculture

June 24 – June 27, 2018

Montreal, Quebec, Canada

Abstract. *The use of computer systems in Precision Agriculture (PA) promotes the processes' automation and its applied tasks, specifically the inspection and analysis of agricultural crops, and guided/autonomous locomotion of mobile robots. In this context, this research aims the application of computer vision techniques for agricultural mobile robot locomotion, settled through an architecture for the acquisition, image processing and analysis, in order to segment, classify and recognize patterns of planting rows, as guiding references for steering the mobile robot. Also, the process includes: filtering in the spatial domain for acquired images; pre-processing in RGB and HSV color spaces; JSEG unsupervised segmentation algorithm, applied to color quantization in non-homogeneous regions; normalization and histograms feature extraction of preprocessed images for training and test sets, fulfilled by the principal components analysis (PCA); pattern recognition and statistical classification. The developed methodology includes sets of 700 and 900 images' databases for each approach of natural scenes under different conditions of acquisition, providing great results on the segmentation algorithm. Statistical classification (Bayes/Naive Bayes) was applied, proving the efficiency in recognizing distinct patterns and classes for images in several characteristics inherent to the robot locomotion environment.*

Keywords. *Agricultural mobile robots, natural scenes recognition, image segmentation, statistical classifiers.*

Introduction

The recent emerging of technologies in embedded electronic systems and automation has promoted the development of scientific research activities directed to Precision Agriculture (AP), focusing on applications of machines and agricultural implements of high performance (PEREIRA, 2009). This development encompasses the application of experimental procedures to several tasks aimed at the application in agricultural mobile robotics:

- computational vision in the application of digital image processing techniques for the localization and analysis of agricultural crops;
- navigation and trajectories' planning for locomotion;
- probabilistic algorithms in graphical information processing and pattern recognition;

In this context, the present research proposes an approach for analysis of planting areas lines in orchards, whose purpose is to delimit the traffic area of a mobile robot, in areas denominated as navigable, based on the angle of guidance and lateral displacement (Lulio, 2011).

Methodology

As methodology, the complexity of the elements discrimination for graphic processing in computer vision demands the following steps:

- preprocessing acquired images for identification of regions of interest (ROI) in RGB/HSV color spaces;
- application of an unsupervised segmentation algorithm proposed in color quantization for nonhomogeneous regions;
- validation of feature vectors, descriptors of patterns in the processed images, based on statistical methods (Bayes/Naive Bayes).

The analysis of image processing depends of the studied intrinsic criteria of orchards, corresponding their physical elements in determining the planting lines for the vehicle navigation. The orchard delimitations of planting area were acquired in frontal perspective vision, which identify possible regions of interest to be segmented as color/texture processes.

Also, the camera distance is determined according to the acquisition, during the robot first locomotion. The acquired images were established at distances of 5m to 0.50m among the planting lines and the robot. Chromatic resolution of these images is 3664 × 2062 pixels / 24-bit color resolution, stored in a JPEG/JPG file, acquired in the RGB color space. A set database for the orchard scenes, of 920 images was applied for all computer vision processing algorithms.

Fig 1. Orchards planting scenes.

The computer vision processing strategy is shown in figure 2.

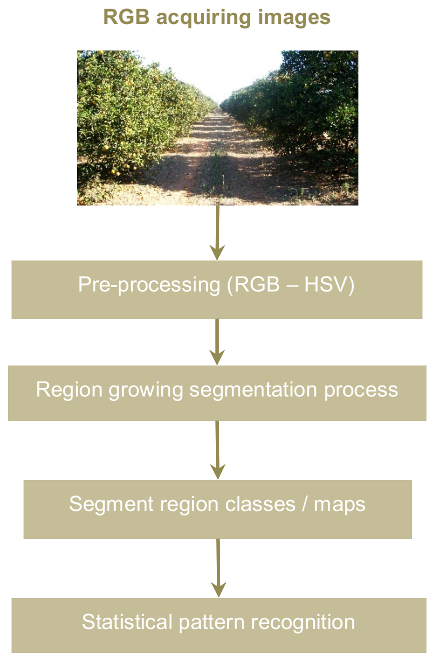

Fig 2. Computer vision processing strategy.

Unsupervised segmentation algorithm based on nonhomogeneous regions

Textured images segmentation considers the spatial grouping of growing pixels in nonhomogeneous regions, converting them onto homogeneous segments, through the relation of different scales in the treated color space. In natural scenes, it is an unsupervised technique for textured and colored regions, which is appropriate for verifying the homogeneity in patterns of regions with these features, and computationally is more feasible and effective than models of parameter estimation.

In this unsupervised segmentation model, homogeneous regions of color/texture, quantization of color information, color distinction among neighboring regions were considered for image analysis. This technique has two processing steps: (a) quantization of the color space (*peer-group filtering – PGF*) (DENG et al., 1999b); (b) growth / grouping of regions with similar colors (DENG et al., 1999a; DENG and MANJUNATH, 1999c; DENG and MANJUNATH, 2001).

In the first step, the perceptual algorithm of quantization is performed in minimal degradation of color homogeneity, defining regions in reduced colors, each associated to a class. The pixels in the original image were replaced by such classes, forming a class map in texture composition. Each class map represents a J-region (*J-image*), denoted by the quantized color area, containing pixels with positive and negative values (*J-values*) to the boundary and the texturing of its boundaries. The growth and grouping of regions with similar colors, as second step, connects the identification of the similarity of colors and the spatial distribution of these colors, delimiting the contour regions of the homogeneity found.

To compute then, the J-value, Z is defined as a set of quantized image points, so $z = (x, y)$ with $z \in Z$, and considering m the mean of all Z elements. C is the number of classes and Z_i are the elements of Z belonging to class i, where $i = 1, ..., C$, and m_i are the means of the elements in Z_i.

$$m = \frac{1}{N}\sum_{z \in Z} z \qquad (01)$$

$$m_i = \frac{1}{N_i}\sum_{z \in Z} z \qquad (02)$$

Therefore, the J-value is given by:

$$J = \frac{S_B}{S_W} = \frac{(S_T - S_W)}{S_W} \qquad (03)$$

Where:

$$S_T = \sum_{z \in Z} \|z - m\|^2$$
$$S_W = \sum_{i=1}^{C}\sum_{z \in Z} \|z - m_i\|^2 \qquad (04)$$

The S_T parameter represents the sum of the quantized points in the image, within a mean in all Z elements. Thus, the relation among S_B and S_W denotes the distance measures of class maps for arbitrary nonlinear distributions.

As seen in figure 3, the algorithm flowchart for spatial distribution is presented.

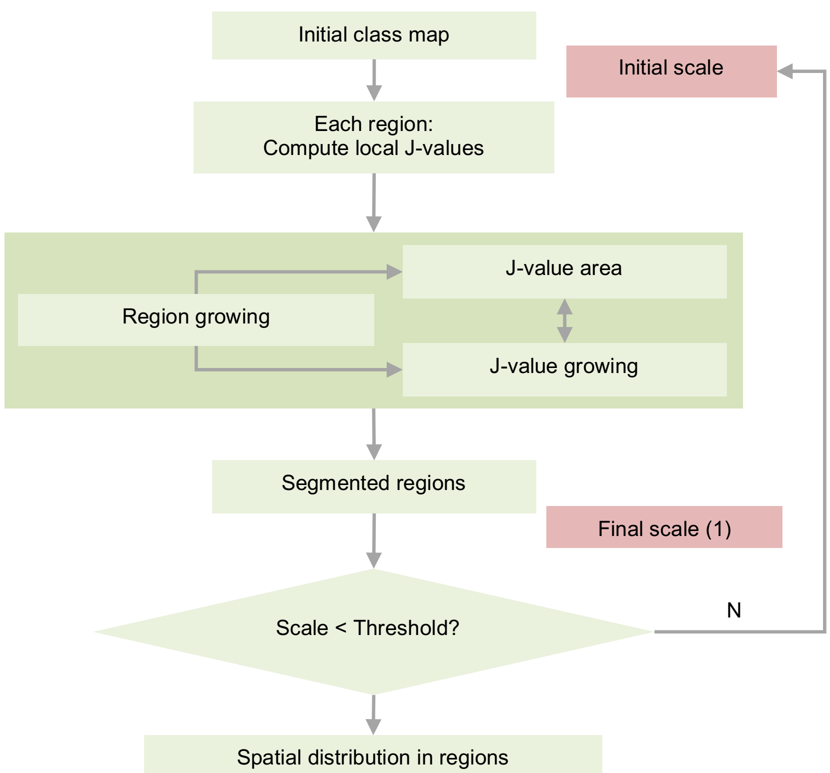

Fig 3. Region growing algorithm.

The initial transformation in HSV color scale provides the attempt to group the segments with greater shade, constituting the first class of pattern (navigable area); the other segments, geometrically allocated in the image, by lateral and upper deviations were related to the second class of pattern (non-navigable area).

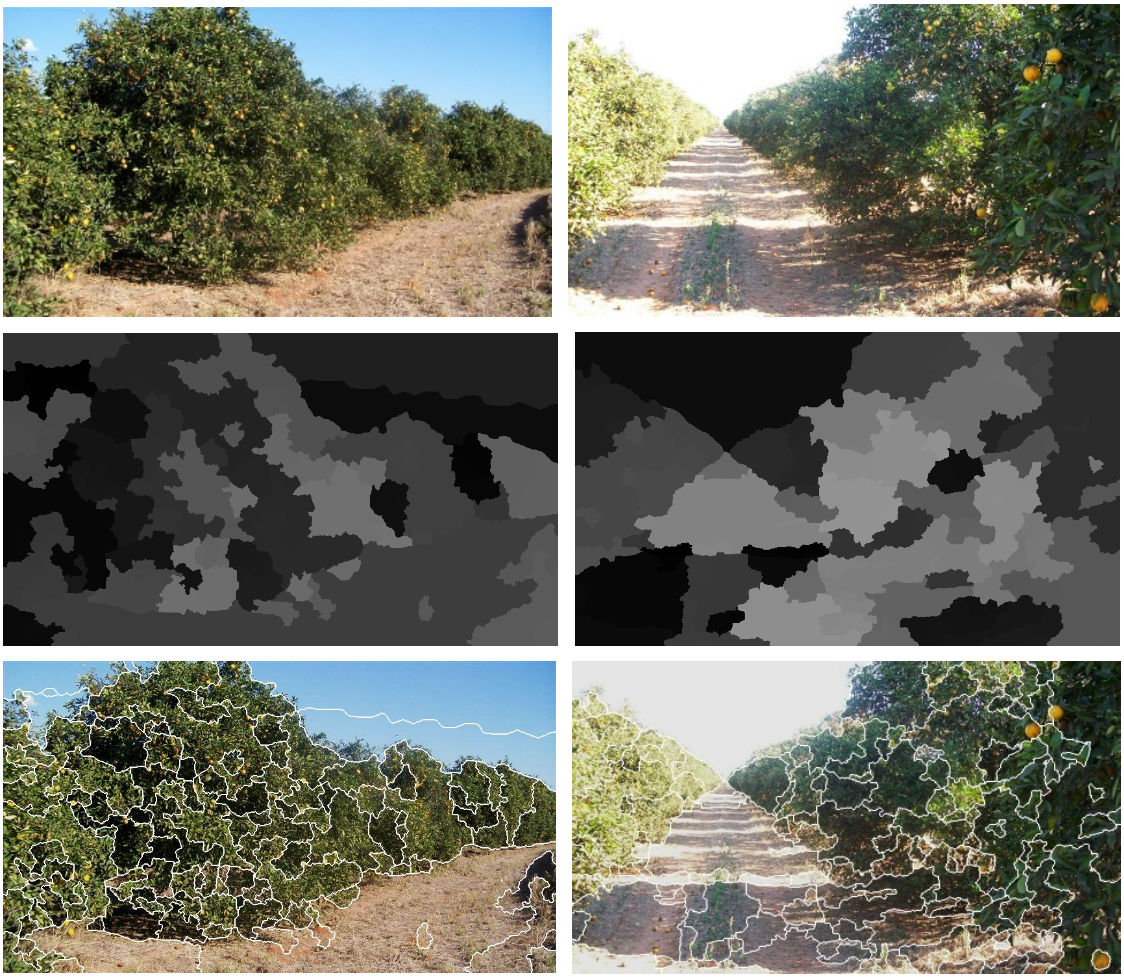

Fig 4. Results of color quantization / spatial distribution for a RGB acquisition.

The distribution of feature vectors to the orchard classes in RGB color space refers to the classes components: *navigable area* and *nonavigable area*, for sets among 10% and 50%. The feature vectors present a majority data in the second class for all percentages, since the images analyzed in this approach present more non-navigable regions, such as the planting lines on both sides and the horizon.

Table 1. Feature vectors distribution for orchards RGB images.

Orchards (RGB)			
%	Navigable area	Nonavigable area	Total
10	3768	4128	7896
20	3876	4024	7900
30	3900	4054	7954
40	3923	4067	7990
50	3941	4154	8095
75	4001	4244	8245
100	4056	4401	8457

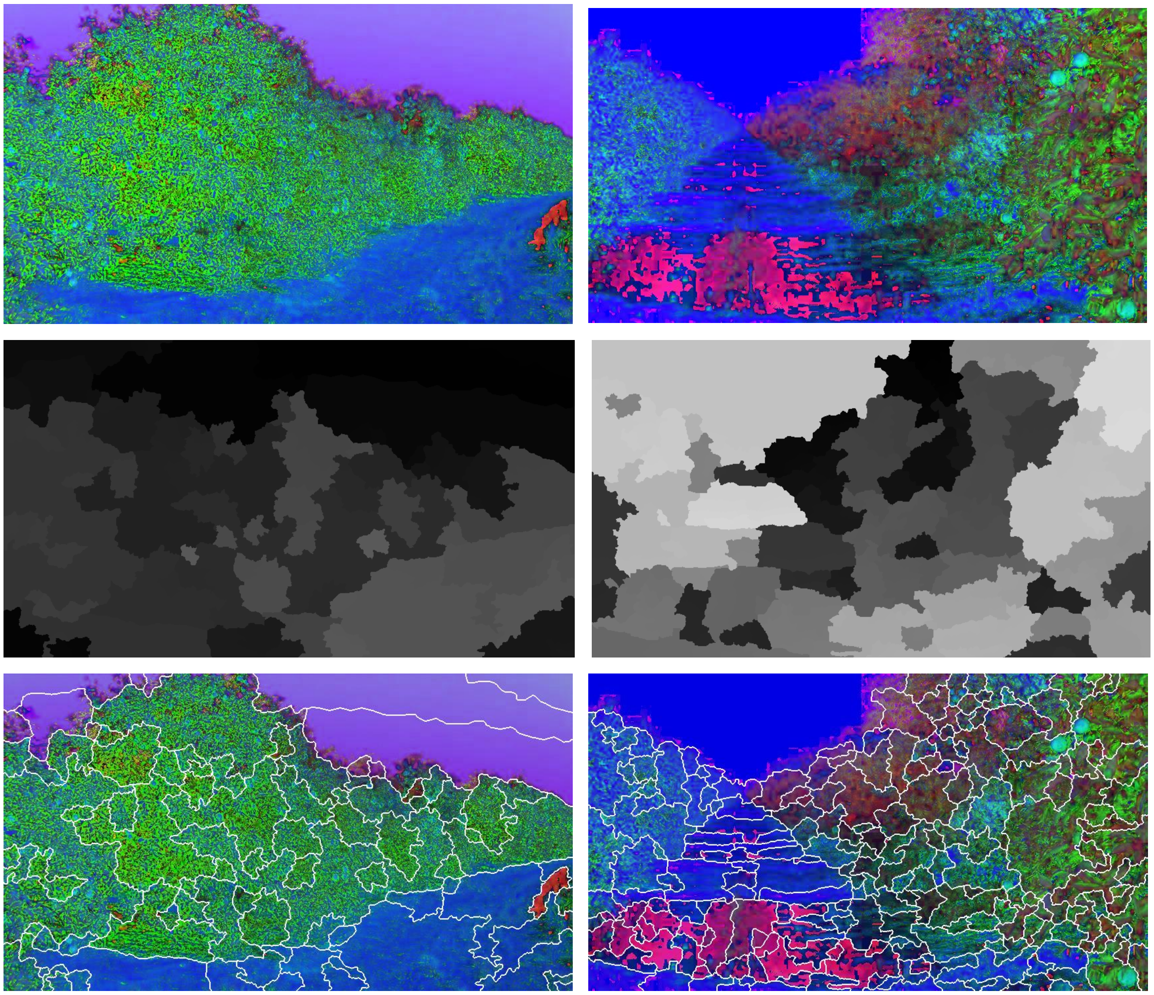

Fig 5. Results of color quantization / spatial distribution for a HSV acquisition.

Similarly, the feature vectors in HSV denote a majority data in *nonavigable area* class due to the same behavior in RGB case, with percentages 75% and 100%, the difference among the total amount of vectors relative to the lower percentages increases, given the concentration of vectors when the region of interest of first class (*navigable area*) is thresholded in the *H*, *S* and *V* components.

Table 2. Feature vectors distribution for orchards HSV images.

Orchards (HSV)			
%	Navigable area	Nonavigable area	Total
10	3770	4132	7902
20	3789	4126	7915
30	3800	4154	7954
40	3832	4197	8029
50	3845	4210	8055
75	3890	4249	8139
100	3916	4414	8330

Statistical methods for pattern recognition

Bayes theorem: Considering the conditional probability is unavailable in cases where a priori probability is used for the estimation of training sets, the Bayes theorem is given by the probability density function (PDF), in equation 05.

$$P(C_i \mid y) = \frac{p(y \mid C_i)P(C_i)}{\sum_{j=1}^{K} p(y \mid C_j)P(C_j)} \tag{05}$$

From eq. 05, *p(y | Ci)* is the conditional PDF of class *i*, where *y* is a scalar or vector value, for several features, whose random variables are denominated multi-variances with joint PDF (PDFc). In this way, the resolution of PDFc depends on the number of components of the mixture (*k*) and its parameters (*q*), which are computed through maximum-likelihood (ML) or Bayesian estimation (PAALANEN et al., 2006).

Therefore, the *expectation-maximization* algorithm for obtaining the ML constrains, estimates through the insensitivity to initialization and the necessity of choosing the number of components for mixture considered as under-adaptation or super-adaptation. The customization of this algorithm was computed by (FIGUEIREDO and JAIN, 2002) that corrects such problems.

The equation for obtaining the PDFc is a model of finite Gaussian function mixtures, where *k* is the number of mixtures, *αm* are the weighted probabilities of the mixture, *θm* – the Gaussian function parameters (FIGUEIREDO and JAIN, 2002).

$$\widehat{p}\left(y \mid \theta\right) = \sum_{m=1}^{k} \alpha_m p\left(y \mid \theta_m\right) \tag{06}$$

The number of components is arbitrarily defined at the initialization of the algorithm, whose means were distributed in the space formed by the training set. These components were eliminated when unsupported by the data, that is, they are saturated. With only the minimum number of components, the algorithm results the most efficient model given the weighting of the objective function, which must be minimized in relation to *θ*.

The estimation of the PDFc is directed to each quantity of dimensions, in all percentage sets. The classification of the set of tests implies in the result of the higher hit rate. The programming of this algorithm is also built/computed as a library to the *Matlab Statistics toolbox*.

Naïve-Bayes theorem: As independent features in this variation of Bayes' theorem, the assumption is given by the degree of uncertainty of the PDF for each class or set of classes.

$$p(y \mid C_i) = \prod_{j=1}^{n} p(y_j \mid C_i) \tag{07}$$

Since *p(y | Ci)* is the conditional marginal density function (PDFc), the training and evaluation of this statistical classifier were made for several features in the discrimination of the PDF, resulting in several discrete marginal PDFs for each dimension analyzed, based on the estimation by maximum-likelihood estimation. This criterion function is given by the mean and variance measures of each class of training standards as data input, given by the normal distribution in the estimation of the PDF of each class.

The graphic 6 illustrates the number of dimensions (feature vectors) in the RGB case, for each class, indicating the maximum quantities relative to each percentage. The two classes approach to their quantities, with fewer dimensions for the first class. The same relation is valid for the HSV case, but the disparity between the two classes is greater.

For HSV processing, the degree of uncertainty in the classification of navigable area is low, since the regions for navigation have low rate of false negatives, unlike what occurs with the RGB color space, in which the non-navigable area class tends to present higher hit rates. The number of dimensions in this case provides better hit rates, with consistent classification of the two classes in HSV and, to a lesser degree, in RGB, so the training set is larger, the prediction requires longer processing time.

The probability of the first class navigable area for RGB case is estimated at a lower index than the second class for all training sets. For HSV case, the same first class presented a higher estimation index than the second class for the training sets at 40%, 50%, 75% and 100%.

Table 3. Probability estimated a priori, for Bayes/Naïve Bayes classifiers.

		Orchards						
Color space	Classes	10%	20%	30%	40%	50%	75%	100%
RGB	Navigable area	0,4129	0,4273	0,3128	0,4112	0,4223	0,4122	0,4187
	Nonavigable area	0,5871	0,5727	0,6872	0,5888	0,5777	0,5878	0,5813
HSV	Navigable area	0,3593	0,4544	0,4384	0,5209	0,5232	0,7733	0,8274
	Nonavigable area	0,6407	0,5456	0,5616	0,4791	0,4768	0,2267	0,1726

Conclusions

Image preprocessing techniques, segmentation, extraction and classification, pattern recognition, and post-processing aggregate the applied procedures of computational vision in this research, with the objectives focused on the autonomous locomotion of mobile agricultural robots.

In the preprocessing, transformations were made in the chromatic resolution and in the color spaces of the acquired images, allowing different alternatives of analysis and segmentation in natural scenes.

The unsupervised segmentation algorithm had an expected result, identifying the homogeneous regions as belonging to the same class, and associating them with the objects of interest in the delimitation of the navigable areas.

In the extraction of feature vectors, dimensionality reduction to empirically acceptable values as input parameters to the training sets was adequately performed by the ACP, providing the necessary information to the classifiers, through the relation among the hit rate and the number of dimensions.

The naive Bayes and Bayes theorem were able to estimate the components of the mixtures among these sets, so that the relationship between the sorted regions and the estimated classes presented viable results for pattern recognition.

References

DENG, Y., KENNEDY, S., MOORE, M. S. e MANJUNATH, B. S. (1999b). Peer group filtering and perceptual color image quantization. Proc. IEEE International Symposium on Circuits and Systems VLSI (ISCAS'99), Orlando, FL, vol. 4, pp. 21-24.

DENG, Y. e MANJUNATH, B. S. (1999c). An efficient low-dimensional color indexing scheme for region based image retrieval. Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP-99), Phoenix, Arizona, March 1999.

DENG, Y. e MANJUNATH, B. S. (2001). Unsupervised segmentation of color-texture regions in images and video. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI '01).

DENG, Y., MANJUNATH, B. S. e SHIN, H. (1999a). Color Image Segmentation. Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR'99, Fort Collins, CO, vol. 2, pp. 446-51.

FIGUEIREDO, M. A. T. e JAIN, A. K. (2002). Unsupervised learning of finite mixture models. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 24, n. 3, p. 381–396. ISSN 0162-8828.

LULIO, L. C. 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, 2011.

PAALANEN, P., KAMARAINEN, J., LLONEN, J. e KALVIAINEN, H. (2006). Feature representation and discrimination based on gaussian mixture model probability densities – practices and algorithms. Pattern Recognition, v. 39, p.1346–1358.

PEREIRA, R. R. D. (2009). Protocolo ISO 11783: Procedimentos para comunicação serial de dados com o Controlador de Tarefas. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo. 188p. São Carlos, SP - Brasil.