The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Pattern Recognition Structured Heuristics Methods for Image Processing in Mobile Robot Navigation

Luciano C. Lulio, Mario L. Tronco, Arthur J. V. Porto

Abstract— **In this project, the main focus is to apply image processing techniques in computer vision through an omnidirectional vision system to agricultural mobile robots (AMR) used for trajectory navigation problems, as well as localization matters. To carry through this task, computational methods based on the JSEG algorithm were used to provide the classification and the characterization of such problems, together with Artificial Neural Networks (ANN) for pattern recognition. Therefore, it was possible to run simulations and carry out analyses of the performance of JSEG image segmentation technique through Matlab/Octave platforms, along with the application of customized Back-propagation algorithm and statistical methods as structured heuristics methods in a Simulink environment. Having the aforementioned procedures been done, it was practicable to classify and also characterize the HSV space color segments, not to mention allow the recognition of patterns in which reasonably accurate results were obtained.**

Index Terms— **pattern recognition, image segmentation, mobile robots, computer vision**

I. Introduction

ALGORITMS for crops inspection and planting area avoidance obstacle navigation, work mainly with images composed of complex objects, textures, shadows and brightness. Several segmentation algorithms proposed in literature [1], [2], [3] were designed to process images originally characterized by the above-mentioned items. Additionally, agricultural automation may take advantage of computer vision resources, which can be applied to a number of different tasks, such as inspection [4], classification of plants [5], [6], [7], estimated production [8], automated collection [9] and guidance of autonomous machines.

Bearing the afore-named in mind, the present project used both JSEG segmentation algorithm [10] and multilayer perceptron (MLP) in order to segment and therefore classify images into the following classes: navigable area, planting area, and sky. The approach tried to segment classification deploys an artificial neural network (ANN) – iRPROP algorithm, faster than usual *Back-propagation* algorithm [11] – to classify and characterize the segments into three classes. A feature vector was formed with color channels histograms. After training, the mean squared error (MSE), which was obtained in the different ANN topologies, indicated the most appropriate topology. Ultimately, the results achieved by segment classification were used to create the image-class map. After segmentation, it was necessary to identify and classify the segments. Each one can be represented by a feature vector. Several metrics (vector bundle) can be part of a feature vector, however, a subset of those which describes and evaluates appropriate classes of segments was chosen.

This article presents the following structure: 1) construction of the system unit – omnidirectional vision system modeling; 2) polar-to-cartesian transformational method; 3) JSEG image segmentation and processing; 4) classification results from probabilistic methods and ANN for pattern recognition matching [12]; 5) conclusion and forthcoming works.

II. Omnidirectional Vision System Modeling

Several previous works have described different kind of mirrors in catadioptric systems for omnidirectional vision acquisition as described below. Typically, these systems consist in a convex mirror fixed in front of a camera that also remains fixed with its lens facing up.

A vision system set with a single center of projection [13] generates any projected image on any defined and known plane. So it is possible to generate perspective or panoramic images without distortions. In more specific works, a vision system is developed, which uses an orthographic projection camera and a parabolic mirror, with simulation software that allows manual adjustment of perspective projection properties (vision virtual direction, field of view and camera magnitude) [14], [15]. Another

L. C. Lulio is with Mechanical Engineering Department – Engineering School of Sao Carlos – University of Sao Paulo, CEP 13566-590 Brazil (phone: 5516-33739247; fax: 5516-33739257; e-mail: lclulio@sc.usp.br).

A. J. V. Porto is with the Mechanical Engineering Department – Engineering School of Sao Carlos – University of Sao Paulo, CEP 13566-590 Brazil (phone: 5516-33739247; fax: 5516-33739257; e-mail: ajvporto@sc.usp.br).

M. L. Tronco is with Department of Computer Science and Statistics – State University of Sao Paulo CEP 15054-000 Brazil. He is also with Mechanical Engineering Department – Engineering School of Sao Carlos – University of Sao Paulo, CEP 13566-590 Brazil (phone: 5516-33739247; fax: 5516-33739257; e-mail: mariot@ibilce.unesp.br).

work [16] uses an omnidirectional system in an application of visual servo control of a mobile robot, thus avoiding that a pan/tilt mechanism point directly to the interest target.

A catadioptric system with single center of projection was also developed, which is composed of a perspective projection camera and a hyperbolic mirror, which allows the acquisition of perspective and panoramic projections without distortion. The author also shows the epipolar curves/geometry for this type of system, which calculates distances based on stereo vision. The epipolar geometry between two images of the same scene, captured by similar cameras, is the intersection geometry of both image planes (I and I'), with the planes bundle that has the base line (line that connects both optical centers, O and O') as axis [17].

Among the above-mentioned referenced works, this one aims to build a hyperbolic mirror similar to the catadioptric system described above. The hyperbolic mirror with a perspective projection camera would allow image acquisition with single projection center in an omnidirectional system (panoramic or perspective images).

Omnidirectional computer vision systems make the acquisition of image within a 360° field of view practicable. The arrangement of convex hyperbolic mirrors with lenses makes a computer vision system structure with larger vertical field of vision [18], [19], [20].

The angle of elevation (φ) is the angle of incidence of light on the mirror surface, and the radial angle (ψ) corresponds to the angle of reflection of light into the image acquiring device. Thus, a pixel coordinate in perspective projection is related to a light-ray direction that reaches the camera sensor and forms a pixel in the polar image, as illustrated in Fig. 1 and Fig. 2.

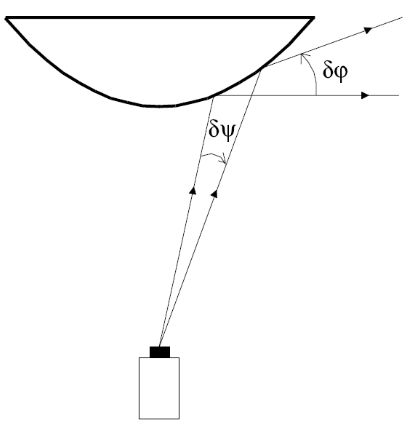

Fig. 1. Vertical angular gain.

The above-named relation can therefore be determined by geometrical functions and exemplified by the following equations.

$$\tan\phi = \frac{f_p sin\phi_o + v_p \cos\phi_0}{f_p \cos\phi_0} \quad (1)$$

$$\tan\theta = \frac{(f_p \cos\phi_0 - v_p sin\phi_0) sin\theta_0 - u_p \cos\theta_0}{(f_p \cos\phi_0 - v_p sin\phi_0)\cos\theta + u_p sin\theta_0} \quad (2)$$

To correct the panoramic images into perspective ones, the method namely direct transformation from Polar to Cartesian coordinates was employed.

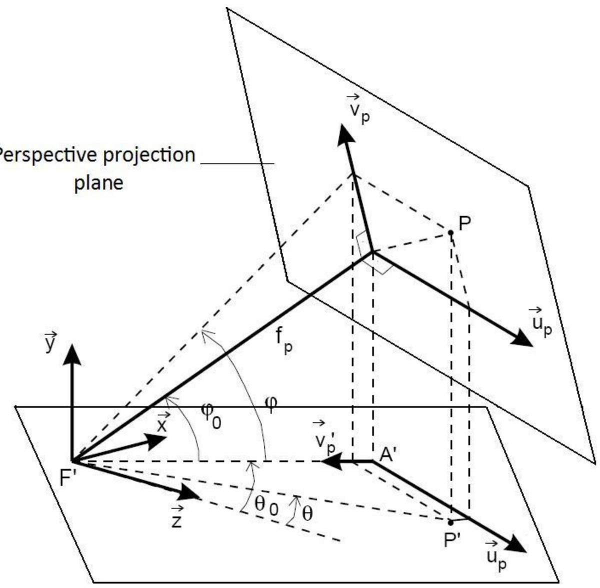

Fig. 2. Geometric relations of the perspective images.

III. Polar-to-Cartesian Transformation Method

This correction method is based on a Polar-to-Cartesian transformation. The radius of the omnidirectional image acquired by the system is linearly mapped as the y coordinate and the angle in the same image are mapped in the omnidirectional x coordinate of the rectified image (x_{in} and y_{in} as input omnidirectional coordinates and x_{out} and y_{out} as output panoramic coordinates). According to this method, if the center of the image is considered the origin (0,0) and all variables assume values between 0 and 1, equation (3) is used to do the correction.

$$\begin{aligned} x_{in} &= y_{out}.\cos(2\pi x_{out}) \\ y_{in} &= y_{out}.sin(2\pi x_{out}) \end{aligned} \quad (3)$$

A panoramic image acquisition in a 360° field of view is transformed from Polar to Cartesian coordinates. The azimuth angle, also known as meridian angle, which manifests itself by the incidence direction of the light-ray onto the angular compass in the coordinated original image, is mapped on the horizontal coordinate axis of the panoramic image, and the radial coordinate in original image is mapped on the vertical coordinate axis of the same image, with omnidirectional, panoramic and binary imaging view processes as illustrated in Fig. 3.

IV. JSEG Image Segmentation

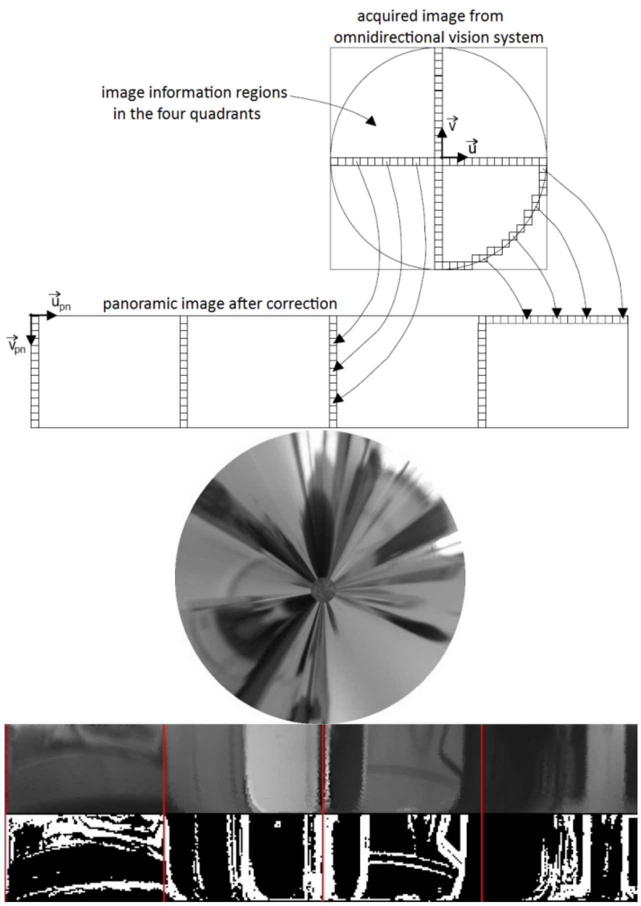

Fig. 3. Omnidirectional, panoramic and binary imaging process after Polar-to-Cartesian correction method.

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) [21], [22]. 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 [23] 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.

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. The J-value is as follows:

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

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

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

The sequential images, on Fig. 4 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.

Fig. 4. JSEG image segmentation process.

V. SEGMENTS CLASSIFICATION AND ANN

Due to the nature of nonlinear vectors, it is fundamental that an ANN-based pattern classification and recognition be used. *Multi-Layer Perceptron* (MLP) [24] was implemented through a customized *back-propagation* algorithm for complex patterns.

Derived from back-propagation, the iRPROP algorithm (improved resilient back-propagation) [25] is both fast and accurate, with easy parameter adjustment. It features an Octave [26] 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.

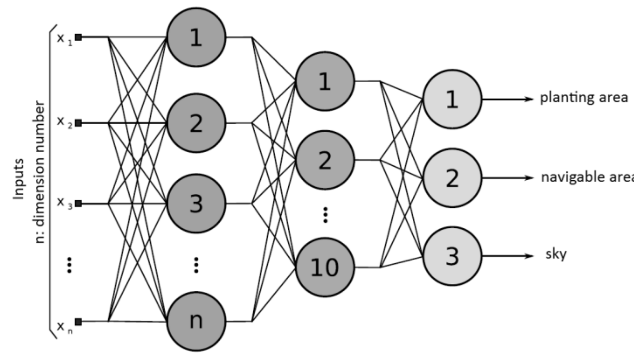

Fig. 5. ANN schematic topology.

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 [27] for Matlab (Mathworks Inc.) platform. The most appropriate segment and topology classifications are those using vectors extracted from HSV color space. Also, a network with less MSE in the HSV-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. Tab. I shows the previously data.

TABLE I
MSE RESULTS FOR EACH TOPOLOGY

MSE	Neurons	Navigation area	Planting area	Sky
HS	32	0,089143	0,094905	**0,023409**
	64	0,099398	0,045956	0,089776
	128	0,049100	0,095064	0,097455
	256	0,057136	0,099843	0,034532
HSV	32	0,089450	0,022453	0,067545
	64	0,059981	**0,010384**	0,082364
	128	0,049677	0,078453	0,043493
	256	**0,038817**	0,079856	0,045643

VI. NORMALIZATION AND FEATURE EXTRACTION

This section tackles how statistical methods were employed as a combination of results with ANN as structured heuristics method, 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 patter classification. The MSE results for each topology, shown in Table 1, were partitioned to eliminate the feature vectors that are distant from the class centroids, so the classifier will deal in less dispersed vectors. Upon observing the following table, which shows the vector distribution in five training sets (20%, 30%, 50%, 70% and 100%), this work approached two probabilistic classification methods in order to match final pattern recognition results with ANN: Bayes theorem and Naive Bayes.

TABLE II
VECTOR DISTRIBUTION FOR RGB AND HSV SPACE COLORS

	RGB			HSV		
%	NA	PA	Sky	NA	PA	Sky
20	1029	5486	34	1024	5384	26
30	1345	5768	54	1342	5390	45
50	1390	6094	130	1390	6003	103
70	1409	6298	149	1402	6209	140
100	1503	6300	158	1402	6209	145

Navigation area = NA; Planting area = PA;

RGB space color is used to compare the total number of dimensions in feature vectors with HSV. With a smaller dimension of iterations, HSV was chosen as the default space color. For such iterations inspection, a technique (main component analysis – MCA) 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 [28], [29], [30].

In HSV space color, the Bayesian classifiers have produced results which are similar to RGB, where there is a hit rate when the number of dimensions increases in an accuracy average ranging from 20% to 50%, as can be seen in Tab. I. A maximum rate accuracy for HSV is 0.38817, which occurs for 30% and 6777 dimensions. In RGB space color, Bayesian conventional classifiers are identical to Naive results, because as the dispersion of classes increases, there is an average hit rate, which goes up to 50%. The classifiers concerning the number of dimensions are different from the previous ones, which range from 20% and 30%, where hit rates fall as the number of dimensions increases.

As a consequence, Bayesian classifiers in HSV space color, outperforms the other classifiers as shown in Fig. 6. The average rate of achievement value, together with the number of dimensions *draw a linear convergence* for all vector distribution in the five training sets. Although the three methods deliver different performances, yet similar behavior, because the hit rate of the class sky tends to improve owing to the increase in the number of dimensions. Navigation and planting area classes are listed as false feature vectors by the texture similarities in training, which means that ANN and Bayesian must be coupled for improved results.

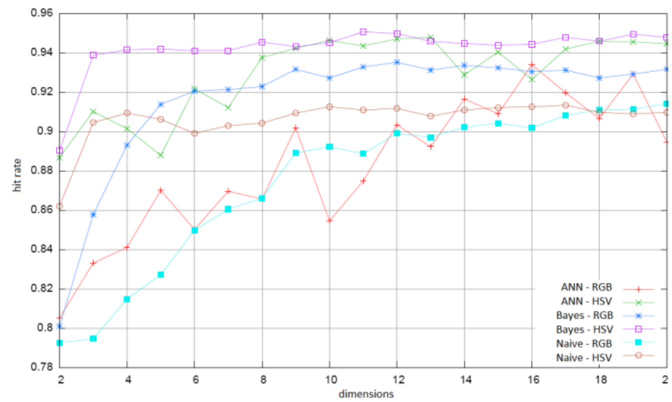

Fig. 6. Average hit rates for the three major training sets of training.

The following graph about the first components shows that the RGB curves have a higher percentage than most HSV curves. Also, it can be observed that all curves present values lower than 90%.

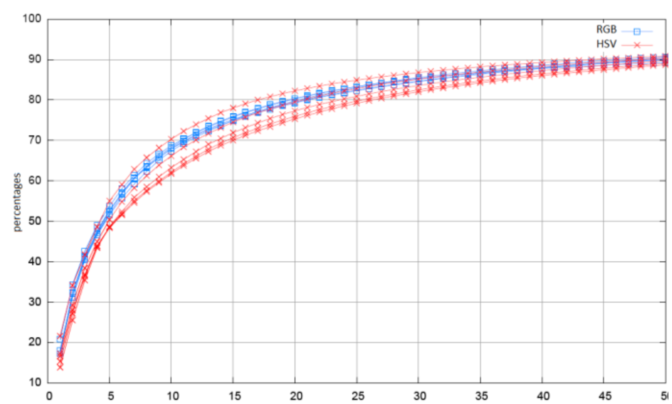

Fig. 7. RGB and HSV relation with amount of dimensions.

VII. CONCLUSIONS AND FORTHCOMING WORKS

In conclusion, this work presented how efficiency was the segmentation and classification of agricultural scenes for navigation problem. As the data provided evince, this generated algorithm fulfils the expectations as far as segmenting is concerned, so that it sorts the appropriate classes. As a result, a modular strategy with ANN and Bayes statistical theorem can be an option for the classification of segments with JSEG algorithm. Both JSEG and MLP proved suitable for the construction of an image recognition system.

Upcoming works are likely to employ other segmentation algorithms such as Gaussian filters and Neuro-Fuzzy system, only for pattern recognition though.

ACKNOWLEDGMENT

The authors would like to thanks to CAPES for funding this research project.

REFERENCES

[1] Y. LIU, X. ZHOU, “Automatic texture segmentation for texture-based image retrieval”. In: *International Multimedia Modelling Conference*, 10., 2004. *Prodeedings*...IEEE Computer Society, 2004. p. 285-290.

[2] J. CHEN, T.N. PAPPAS, A. MOJSILOVIC, B. ROGOWITZ, “Adaptive image segmentation based on color and texture”. In: *International Conference on Image Processing*, 2002. v. 3, p. 777-780.

[3] P.R. HILL, “Wavelet based texture analysis and segmentation for image retrieval and fusion”. 2002. PhD (Thesis) – *University of Bristol, Department of Electrical and Electronic Engineering.*

[4] T. BROSNAN, D. SUN, Inspection and grading of agricultural and food products by computer vision systems – a review. *Computer and Electronics in Agriculture*, v. 36, p. 192-213, 2002.

[5] L. TANG, L. TIAN, B. L. STEWARD, Classification of broadleaf and grass weeds using gabor wavelets and an artificial neural network. *Transactions of the ASAE*, v. 46, p. 1247-1254, 2003.

[6] J. C. NETO, G. E. MEYER, D. D. JONES, A. J. SURKAN, Adaptive image segmentation using a fuzzy neural network and genetic algorithm for weed detection. In: ASAE ANNUAL INTERNATIONAL MEETING, 2003.

[7] B. L. STEWARD, L. F. TIAN, D. NETTLETON, L. TANG, Reduced-dimension clustering for vegetation segmentation. *Transactions of the ASAE*, v. 47, p. 609-616, 2004.

[8] P. ANNAMALAI, W. S. LEE, T. F. BURKS, Color vision system for estimating citrus yield in real-time. In: CSAE ANNUAL INTERNATIONAL MEETING, 2004.

[9] A. PLEBE, G. GRASSO, Localization of spherical fruits for robotic harvesting. *Machine Vision and Applications*, v. 13, p. 70-79, 2001.

[10] Y. DENG, C. KENNEY, M.S. MOORE, B.S. MANJUNATH, “Peer group filtering and perceptual color image quantization”. In: *Proceedings of the 1999 IEEE International Symposium on Circuits and Systems*, 1999a. v. 4, p. 21-25.

[11] C. IGEL, M. HÜSKEN, “Empirical evaluation of the improved Rprop learning algorithm”. *Neurocomputing*, v. 50, p. 105-123, 2003.

[12] L.L. SILVA, “Environments Classification and Caracterization for maps-based Mobile Robot Navigation”. *Masters Degree Dissertation.* DCCE/IBILCE/UNESP, 2008.

[13] S. BAKER, S. K. NAYAR, 1998, “A Theory of Catadioptric Image Formation”, Proceedings of the 6th International Conference on Computer Vision, pp. 35-42, Bombay, India.

[14] S. K. NAYAR, 1997, “Catadioptric Omnidirectional Camera”, Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 482-488.

[15] V. N. PERI, S. K. NAYAR, 1997, “Generation of Perspective and Panoramic Video from Omnidirectional Video”, Proc. of DARPA Image Understanding Workshop, pp. 243-245.

[16] P. CHENG, M. HERBERT, 1998, “Omni-directional Visual Servoing for Human-Robot interaction”, Proc. of IEEE/RSJ International Conference on Intelligent Robotic Systems (IROS),Victoria, B.C., Canada.

[17] T. SVOBODA, T. PAJDLA, V. HLAVAC, 1997, “Central Panoramic Cameras: Geometry and Design”, Research Report K335/97/147, Czech Technical University, Faculty of Electrical Engineering, Center for Machine Perception.

[18] J. GASPAR, N. WINTERS, J. SANTOS-VICTOR, “Vision-based Navigation and Environmental Representations with an Omni-directional Camera. *IEEE Transactions on Robotics and Automation*, v.16, n.6, p.890 -898, 2000.

[19] Y. YAGI, “Omnidirectional Sensing and Its Applications”. *IEICE Transactions on Information and Systems*, v.E82-D, n.3, p.568-579, 1999.

[20] V. GRASSI, “Omnidirectional Vison System for Mobile Robots control”, *Masters Degree Dissertation*. Polythecnic School of University of São Paulo, 2002.

[21] J. MOREIRA, L. COSTA, “Neural based color image segmentation and classification using self-organizing maps”, *In: SIBGRAPI*, 9., 1996. Brazillian Computer Society, 1996. p. 47-54.

[22] L.F. TIAN, D.C. SLAUGHTER, “Environmentally adaptive segmentation algorithm for outdoor image segmentation”. *Computers and Eletronics in Agriculture*, v. 21, p. 153, 1998.

[23] Y. DENG, B.S. MANJUNATH, H. SHIN, “Color image segmentation”. In: *Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, 1999b. p. 446-451, v. 2.

[24] S. HAYKIN, “Neural Networks: Principles and Practice”. Porto Alegre: Bookman, 2001.

[25] F. A. CAVANI, “Omnidirectional Vision System for Mobile Robots”, *Computing Sciences Final Project Course*. DCCE/IBILCE/UNESP, 2004.

[26] J.W. EATON *et* al. *“*Octave”. Avaliable at: http://www.octave.org. (02/12/2006).

[27] S. NISSEN *et al. “*Fann: fast artificial neural network library”. Avaliable at: http://leenissen.dk/fann/ (02/12/2006).

[28] L. F. COSTA, R. M. CESAR JUNIOR, “Shape analysis and classification – Theory and Practice”. 1. ed. Boca Raton, Florida, EUA: CRC Press LLC. ISBN 0-8493-3493-4, 2001.

[29] S. HAYKIN, “Neural networks: a comprehensive foundation”. 2. ed. New Jersey, EUA: Prentice-Hall. ISBN 0-13-273350-1, 1999.

[30] D. COMANICIU, P. MEER, “Robust analysis of feature spaces: color image segmentation”. In: *Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, 1997.