ISSN 2176-5480

22nd International Congress of Mechanical Engineering (COBEM 2013)
November 3-7, 2013, Ribeirão Preto, SP, Brazil

MOBILE ROBOT NAVIGATION IN AGRICULTURAL ENVIRONMENT THROUGH PARALLEL PROCESSING OF PANORAMIC AND PERSPECTIVE IMAGES IN AN OMNIDIRECTIONAL VISION SYSTEM

Victor de Assis Rodrigues[1]
Vinicius Bergoli Trentini[2]
Faculdade de Ciências da Universidade Estadual Paulista "Júlio de Mesquita Filho"
Rua Cristóvão Colombo, 2265
15054-000 São José do Rio Preto, SP
victorassisrd@gmail.com[1]
vinicius_trentini@hotmail.com[2]

Luciano Cássio Lulio[3]
Mário Luiz Tronco[4]
Arthur José Vieira Porto[5]
Escola de Engenharia de São Carlos da Universidade de São Paulo
Av. Trabalhador São-carlense, 400 CP. 359
13560-970 São Carlos, SP
lclulio@sc.usp.br[3]
mltronco@sc.usp.br[4]

Abstract. *Catadioptric Vision Systems have wide vision field of 360-degree of the environment explored, using combination of lenses and mirrors. It is commonly used in structured indoor environments with static obstacles and when submitted in applications operating in outdoor(and rural) environments, encourage arbitrary and various issues that affect the processing of the acquired images. Given this premise, this paper presents the application of a catadioptric omnidirectional vision system applied in agricultural robot navigation between planting rows of orange groves, where obstacles and irregularities of the ground during locomotion are critical points necessary to correct the pre-processing and segmentation algorithms.*

Keywords: *ominidirectional vision, agricultural mobile robots, catadioptric systems, JSEG segmentation*

1. INTRODUCTION

The advances of the past decades in Mobile Robots enable achievement of increasingly complex tasks and require a high amount of information in the environment. In this context, the vision sensors are widely spread tools and able to capture a large set of data from the environment without the need for physical contact, and with very reasonable cost of implementation. It is also possible to observe that in many problems faced in this area, it is interesting to expand the field of view sensors such as the efficiency of the robot is directly linked to this competence has to collect useful information on the environment and to work according to this information.

According to Nayar and Baker (1997) Omnidirectional Vision is responsible for providing techniques to expand the field of view, thereby obtaining an image of the environment in 360°. There are four ways to achieve an omnidirectional vision: using the traditional system of images, rotating the imaging system, using fish-eye lenses, and through catadioptric systems (Nayar and Baker, 1997).

For this work, the catadioptric systems were adopted, as a combination of lenses and mirrors. The advantage of this system is the low cost and low processing time required to obtain complete omnidirectional image. The main disadvantage when compared with other techniques, is the resolution of the captured images, but this condition is directly related to the quality of the camera and the characteristics of the mirror used. Robot navigation (Ericson and Astrand, 2010; Fiala and Basu, 2002; Li et. al., 2011), surveillance (Zhu et. al., 2000) and reconstruction of 3D environments (Scaramuzza, 2008) are examples of areas that these systems can to work.

Among the many profiles of mirrors that can be used in systems reflectors, the hyperbolic profile was chosen, as it provides a good distribution in the reflected image resolution (Svoboda, 1997), and also for having a single center of projection, feature that ensures a image captured from mirror without of distortions.

However, many obstacles arise when these techniques are used in agricultural environments outdoors. Aside from the weather, inclination of the ground, extensive areas for operation, appearance of vegetation, ambient lighting and presence of obstacles dynamics are just examples of frequently encountered problems (Li et. al., 2009). Moreover, the techniques used must be efficient and applicable at runtime, because this condition is very important in algorithms for the navigation of autonomous robots.

The present work presents a study on segmentation techniques applicable Omnidirectional catadioptric images. The aim here is to find regions in navigable plantations orange groves sufficient for Mobile Robots can operate. This is handled as an operation of a set of selected filters for processing and segmentation algorithm (JSEG), applied in two groups of pictures: in omnidirectional images were directly recovered by the reflection mirror and panoramic images generated from the transposition of coordinates based on the original omnidirectional images. For the cartesian coordinates of these images, the single center of projection method is applied - feature to be explored in hyperbolic mirrors.

About the organization of this work: Session 2 introduces the fundamentals of the method used for structuring research. The Section 3 describes the steps that were performed to create the experimental test and Section 4 presents the obtained results.

2. METHODS AND TECHNIQUES

2.1 Omnidirectional Image Unwrapping

According to the work of (Grassi Jr. and Okamoto, 2007) and (Lulio, L.C., *et* al., 2010), hyperbolic profile mirrors, when used in conjunction with perspective projection cameras, feature a single center of projection. This feature ensures the creation of panoramic images or horizontal perspective without distortions since the light rays that form the image pass through the focal center of the mirror after they were reflected by the surface. This ensures that a single light ray is not reflected more of a time on the same surface region, thus avoiding duplicate images.

For the planning of omnidirectional images, also known as rectification process, Grassi Jr. and Okamoto (2007) present two methods that can be used for the purpose of seizing the property single center of projection. The first, called "Image Unwarping Using the Single Centre of Projection", seeks to create panoramic images by projecting the *pixel*s that form the omnidirectional image in a cylinder around the vision system, since the Center Unified Projection allows a full view of environment through a single point. For the correspondence of image *pixel*s on omnidirectional rectified image, are used Eq. 1, Eq. 2 and Eq. 3:

$$r_i = \frac{x \cdot D \cdot V_{pn} \cdot r_{pixel} \cdot (2e + y_{top})}{(D \cdot V_{pn} - v_{pn}) \cdot x \cdot y_{top} + D \cdot V_{pn} \cdot 2e \cdot r_{top}} \quad (1)$$

$$u = r_i \cdot \cos\left(\frac{2\pi \cdot u_{pn}}{H_{pn}}\right) \quad (2)$$

$$v = r_i \cdot \sin\left(\frac{2\pi \cdot u_{pn}}{H_{pn}}\right) \quad (3)$$

where **u** and **v** represent the coordinates of a *pixel* in the omnidirectional image, $\mathbf{u_{pn}}$ and $\mathbf{v_{pn}}$ represent the coordinates of a *pixel* panoramic image and $\mathbf{H_{pn}}$ represents the length of the panoramic image. To get length, which represents the radial distance of a point in the omnidirectional image, it is used the value of **x** as the radial coordinate of a point on the surface of the mirror and $\mathbf{D.V_{pn}}$ representing the horizontal size of the picture with respect to the mirror surface. Some other parameters are known *a priori*, as **r** *pixel* being the size of the radius grinding, **2e** eccentricity of the mirror, $\mathbf{r_{TOP}}$ the **x** coordinate of the top mirror and $\mathbf{y_{TOP}}$ the **y** coordinate at the top of the mirror.

The second method presented in the work of Grassi Jr. and Okamoto (2007) is called "Omnidirectional Image Unwarping for Generating Perspective Images" and the main idea is to create perspective images based on the projection of a plane in space that is perpendicular to a line passing through the focal center of the mirror. In this case the Eq. 4 is used, Eq. 5 and Eq. 6 for the correlation of coordinates:

$$\tan\varphi = \frac{f_p \cdot sen\varphi_0 + v_p \cdot \cos\varphi_0}{f_p \cdot \cos\varphi_0} \quad (4)$$

$$u = \frac{x \cdot (2e + y_{top}) \cdot r_{pixel}}{(x \cdot \tan\varphi + 2e)} \cos\theta \quad (5)$$

ISSN 2176-5480

$$v = \frac{x \cdot (2e + y_{top}) \cdot r_{pixel}}{(x \cdot \tan\varphi + 2e)} \sin\theta \tag{6}$$

where φ represents the projection angle of the plane, θ represents the vertical angle of the vision system and f_p represents the zoom of image perspective. The other parameters are present in the equations of the first unwrap method and have already been addressed.

Depending of the hardware used, the unwrapped process can be slow and not used in applications requiring real-time response. To assist in this task, Grassi Jr. and Okamoto (2007) suggest the use of lookup tables. Thus, the coordinates of the *pixel*s have been calculated before and stored in a data structure access (in this case, an array). As the image data in OpenCV also corresponds to a array at the time of grinding was required only go two vectors and update the corresponding *pixel*s.

2.2 JSEG Algorithm

For analysis of natural scenes, it is necessary that the chosen segmentation algorithms take into consideration also colors and textures of objects in the images, since this environment regions are formed by non-homogeneous color. The JSEG is a segmentation algorithm originally proposed by Deng, et. al. (1999) and seeks to divide the image into heterogeneous regions using two independent processing stages: color space quantization and spatial segmentation.

According to Yuanjie, et. al. (2006), good results with JSEG are acquired when images have high spatial information of color distribution. For this segmentation, the JSEG algorithm has three steps: color quantization, region growing and merging of regions with similar colors.

In the step of color quantization, a small percentage is calculated degradation in color space using the algorithm of quantization (Deng, et. Al., 1999) with minimal staining. Thus, it is possible to associate each color with a respective class and to map the image where each *pixel* is replaced by the respective class. At the end of this process one obtains a map of classes that will be used in the next step of the algorithm.

After this process and even before the second stage, it is necessary to create a J-image (Lulio, L.C., 2011), where data will be used as a similarity criterion for region growing algorithm. For this, a window is positioned on the quantized image and the *pixel* in the center of the window defines the J values, which makes up the *pixel*s of the J-image. To calculate the value of J, Z is defined as the set of all points in the image quantized and becomes $z = (x, y)$ with $z \in Z$. From this result is obtained m being the average of all values belonging to Z and C that is considered as the number of classes obtained in quantization. Z_i is defined as the elements belonging to the class C_i, where $i = 1, ..., c$ The value J is defined as the following Eq. 7, Eq. 8 e Eq. 9:

$$J = \frac{S_B}{S_W} = \frac{(S_T - S_W)}{S_W} \tag{7}$$

$$S_T = \sum_{z \in Z} \|z - m\|^2 \tag{8}$$

$$S_W = \sum_{i=1}^{C} \sum_{z \in Z} \|z - m_i\|^2 \tag{9}$$

Besides the color quantization, the sequences of images also represent space segmentation as edges and regions with textures.

Different window sizes are used by the J-values: The limits wider detect parameters associated with the texture and the smallest detect changes in color and intensity. For each window size is associated with a scale image analysis. Joining the J-image with different scales is possible to obtain a segmentation of regions with reference to the texture parameters.

The regions with few values of J-image are called valleys. The lowest values are applied with a heuristic algorithm. With this it is possible to determine the starting point of an efficient growth, results obtained by the association of similar valleys. When there are no more *pixel*s to be aggregated in the regions, the algorithm ends.

ISSN 2176-5480

3. EXPERIMENTS

3.1 Softwares andequipments used

For the experiments of this work, a *Logitech* HD C525 camera is used. The vision system was built using a glass cylinder with dimensions 245mm x 79.1 mm x 79.1 mm. The bases were made of acrylic, where the lower base was adapted a stand with a height of 12 cm made of MDF (medium-density fiberboard) for allocating the camera and the top base was allocated the hyperbolic mirror.

A task quite exhaustive in the construction of the system was to correctly centralize the omnidirectional camera with focal center of the mirror. Despite measures are known *a priori* what appeared to be the trivial task complexities shown, since during handling of the parts is virtually inevitable that the inclination and height of the camera suffer small alterations, making this system very sensitive. Once made, this structure was allocated on the back of a motorized quadricycle, for terrestrial mobile agricultural inspections, from "Núcleo de Ensino e Pesquisa em Automação e Simulação" (NEPAS) of "Departamento Engenharia Mecânica da Escola de Engenharia de São Carlos" (EESC-USP), for purposes of agricultural research. A MDF structure was used to raise the position of the support with respect to ground and thus get better results. In Figure 1 is presented the holder for experiment (a) and the allocation of this structure on the vehicle (b).

(a) (b)

Figure 1. Representation of the structure of omnidirectional vision system (a) and location of the camera in quadricycle (b).

Finally, for algorithms programming, the OpenCV library was employed, which provides many resources and implementations techniques for recognizing in the area of computer vision, such as frequency filters, edge detection, tracking and segmentation. Together we used Microsoft Visual Studio 2008 IDE and all codes were written in C++ programming language. The machine used has Intel Core i5, 4GB of RAM and Microsoft Windows 8 operating system.

3.2 Capture of Omnidirectional Images

The images of natural scenes were captured on a farm of Embrapa (Brazilian Agricultural Research Corporation) located in the city of São Carlos/SP. The OpenCV library allows reading and video recording in AVI format, and from that mechanism was possible to capture frames and process them.

The resolution of the images obtained by the vision system was 640x480 *pixel*s. To optimize the results obtained by segmentation techniques, a mask was used to crop the image in order to leave only represented the information about the environment, obtained by reflection of the mirrors.

In the experiment, the vehicle was driven between the planting rows of orange groves by a human operator. After the capture step, we applied computational techniques that are described in Section 3.3. Some problems related to environment aspects and also the building structure when in movement are described in Section 3.4.

To assist with some problems reported in this study, all images were captured with auto-focus mechanisms and adjust lighting activated. The auto-focus allows a calibration of the camera be automatically, seeking a better positioning of the lenses according to the distance to the objects of interest are getting a picture clarity. During testing it was possible to obtain better results with this property enabled even if the mirror remains motionless during the trip. With respect to the adjustment of lighting, it is necessary to enable this property to get better quality images in outdoor environment where lighting may vary at any time.

ISSN 2176-5480

3.3 Image Processing Techniques

According to Grassi Jr. and Okamoto (2007), the unwrap step in omnidirectional systems is important in many applications because the planned image obtained at the end of this process is more easily understood by humans, and also allows them to be applied by definition many algorithms that work only in panoramic images or perspectives. For autonomous navigation, depending on the requirements of the problem, this step may become unnecessary and it is assumed that should be avoided due to the additional processing required. For this reason, the image processing techniques used were defined by applying the priority in omnidirectional images and only later verified the efficiency of this method in the unwrapped images.

The JSEG algorithm applied is described above. As known in literature (Lulio, L. C., 2011), the JSEG is not a technique for real-time navigation because it has a processing latency which prevents the application in mobile robot with immediate response, but can be used for environment mapping. The idea of this paper is to perform a frontal navigation using stereo vision, described in Trentini et. al. (2013), and at this stage to capture the images in the environment with the omnidirectional vision, process them, and thus identify possible regions of navigation. In this case the omnidirectional vision can aid in navigation's maneuver in any direction made by the quadricycle and can also be used for other utilities, such as estimated production, environment mapping and detection of obstacles.

3.4 Environmental problems

As described earlier, the agricultural environments have unique features compared to other environments, and these difficult conditions compromise the quality of images obtained by a fragile system structure, such as omnidirectional catadioptric one. An element of the environment present in many captured images is the sun, as illustrated in Fig. 2 (a). Even with the camera having an automatic brightness adjustment, some frames were heavily influenced by the element. Another problem was related to tarnishing of structure, as over time the glass accumulated water droplets as a result of the temperature inside the enclosure caused by the condensation of humidity and sun exposure.

Another problem was related to uneven ground. The oscillation caused along with the degree of stability of the quadricycle is a factor when the camera is not completely flat in support. Particularly in this work, distortion occurred due to use of an inappropriate material for internal support allocation of the camera. The distortions, as illustrated in Figure 2 (b), not only hinder the characterization of the navigation ways as also possible to be even on the unwrap, because frequently the center of omnidirectional images (which should be the location of the camera in the image) was not in the center of the acquired image. As the grinding process has a pivotal location of the center of the omnidirectional image, it is necessary to implement a routine to find the center of the image before rectification. By default, the center identified omnidirectional image as the position the camera is found on this image. For this, the *Match Template* function, available in OpenCV, was applied, which assured us this process with a time quite efficient.

(a) (b)

Figure 2. Presence of element sun in omnidirectional images (a)
and oscillation caused by uneven ground (b).

In addition to the uneven ground, other problems of the physical environment are the occlusions caused by obstacles and also the massive amount of shadows in the environment. In Figure 2 (a) may be observed a person (obstacle) in the region to the left in the central height of the image. In this case, the segmentation algorithm should highlight that person separately from the navigation region. Another case shown in Fig. 3 (a) there is the occlusion caused by the sheets. This situation is more critical for the navigation algorithm, because it leaves the separating regions of the remaining data is insufficient for correct identification of navigable areas.

ISSN 2176-5480

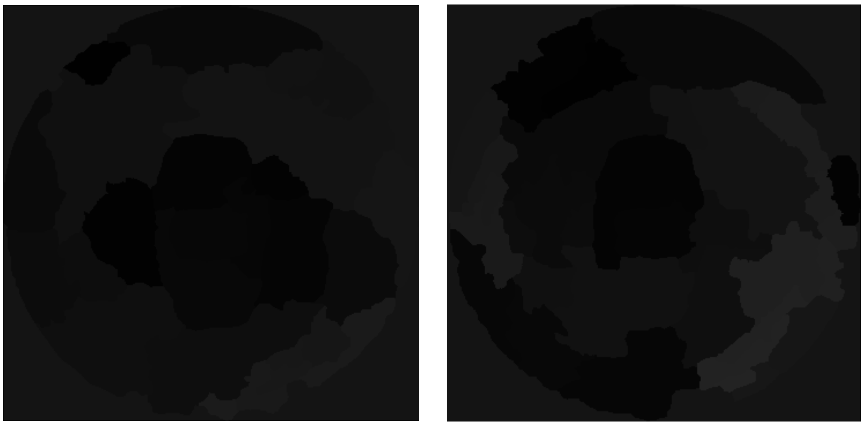
Another condition found is the lack of vegetation in some parts of the plantation regions. In some cases, as in the case of Fig. 3 (b), and the scarcity of plants in border regions of navigation ways is also noticeable change in color of the ground. These changing characteristics render the application of techniques such as threshold, since this value be adjusted.

(a) (b)

Figure 3. Occlusion caused by the sheets (a) and scarcity of plants in border regions (b).

4. EXPERIMENTAL RESULTS

For testing images, 2364 scenes were obtained and processed. The input quantization threshold was set for 250 and 0.1 for merging regions. As a result, this algorithm returns a grayscale image showing the segmented regions. Figure 4 shows an example of the result of excess shadows conditions (a) and occlusions caused by leaves (b).

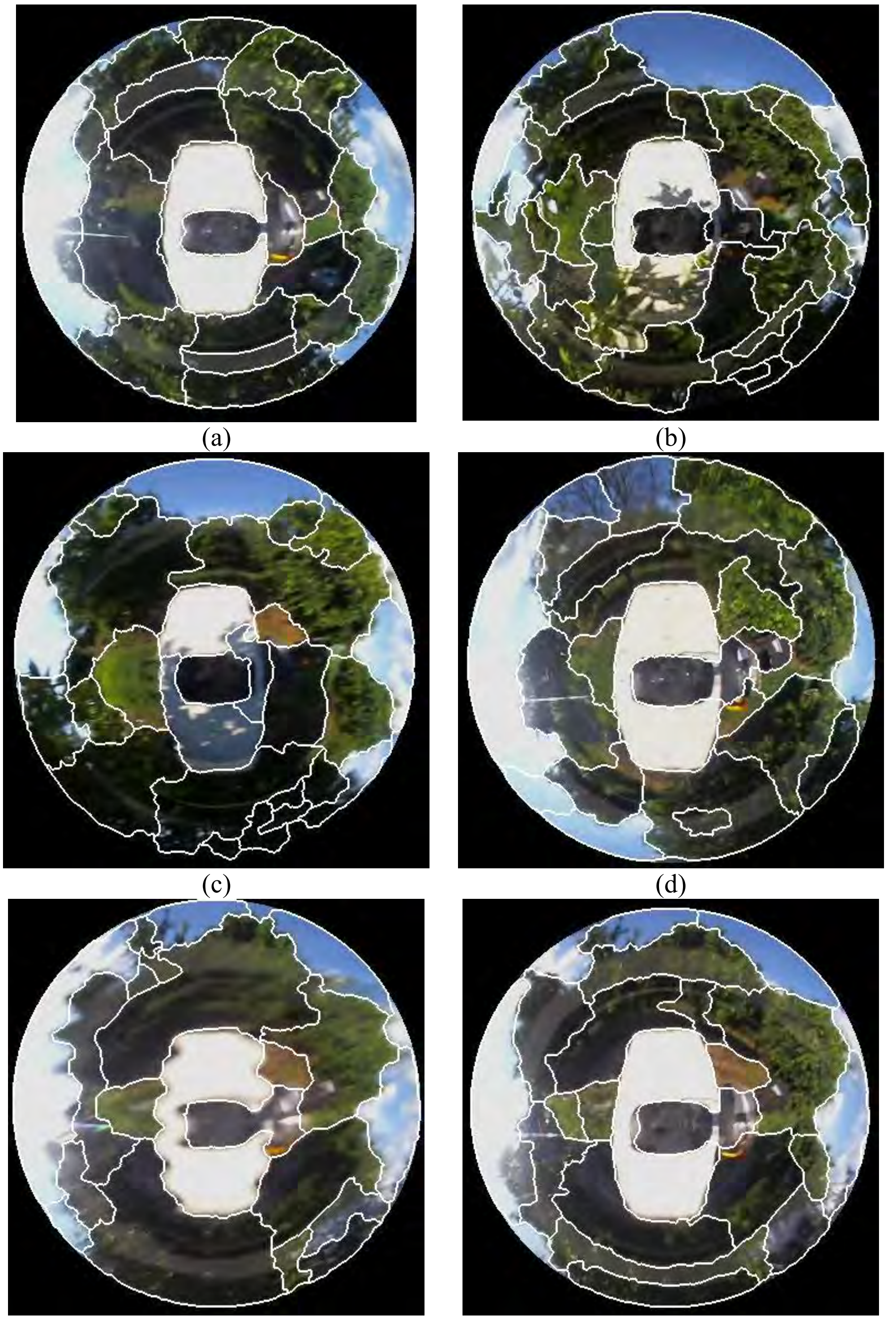


(a) (b)

Figure 4. Grayscale image representing the results under the conditions of excess shadows (a) and occlusions caused by leaves (b).

Figure 5 shows some of the results obtained from the omnidirectional images representing various scenarios and conditions, as a minimum level of interference element and also sun shades (a) caused by occlusions sheets (b) excess of shades (c) little vegetation (d), image distortion caused by the irregularity of the terrain (e) changes in the pattern and texture of the ground (f).

ISSN 2176-5480



(a) (b) (c) (d) (e) (f)

Figure 5. Result JSEG in various situations of natural scenes such as high light intensity (a) occlusions caused by leaves (b), over-shadows (c), little vegetation (d), image distortion caused by the irregularity of the terrain (e) changes in the pattern and texture of the ground (f)

Panoramic images were generated using the method described in Section 2.1 and the same values were applied for JSEG algorithm parameters. Figure 6 shows an example of a unwrapped image (a) and the result of color quantization (b).

(a)

ISSN 2176-5480

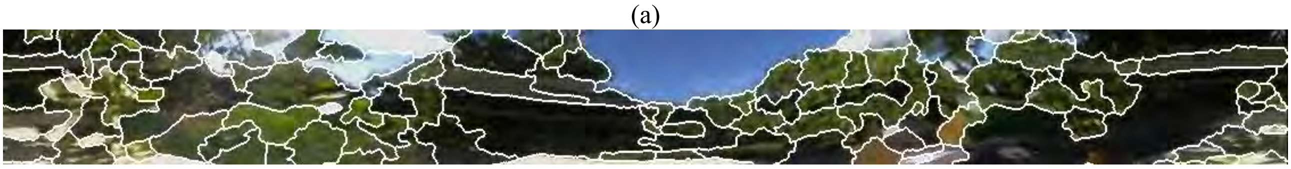
(b)

Figure 6. Panoramic image after rectification step (a)
JSEG color quantization step (b).

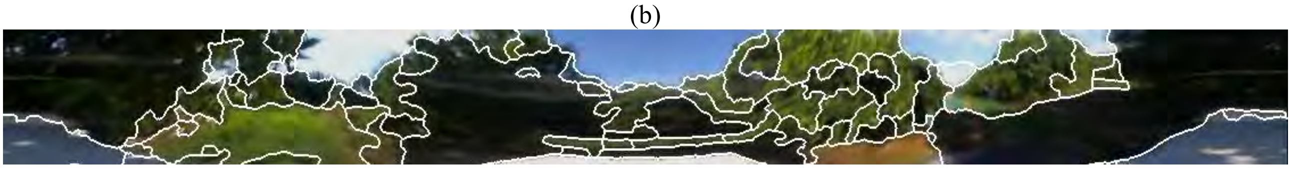
As rotating the image, an overview of navigable lines come to the center of the image and not laterally, but keep the position as shown in Fig 6 to maintain complete visualization of the views front and rear of the vehicle.

Figure 7 shows the results obtained in these scenarios of Fig. 5, in which the interference element and also sun shades (a) caused by occlusions sheets (b) excess of shades (c) little vegetation (d) distortion image caused by the irregularity of the terrain (e) changes in the pattern and texture of the ground (f).

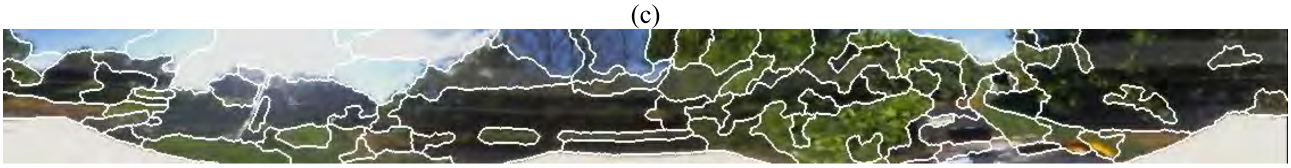
(a)



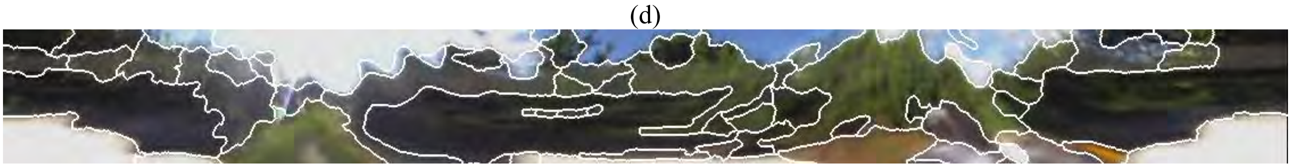
(b)



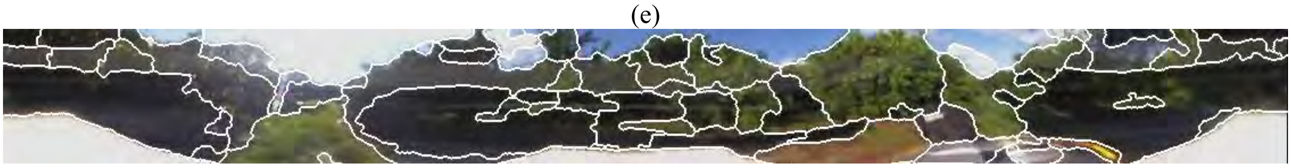
(c)



(d)



(e)



(f)

Figure 8. Result JSEG step after rectification in various situations such as high light intensity (a)
occlusions caused by leaves (b), over-shadows (c), little vegetation (d), image distortion
caused by the irregularity of the terrain (e) changes in the pattern and texture of the ground (f)

5. CONCLUSIONS

In this work it was possible to segment and process images in unstructured outdoor environments suitable, even with several difficulties using omnidirectional catadioptric system in agricultural scenes. Because the attributes of the environment, the physical system construction ends up being the most critical part of the project and many difficulties appear only in the testing step.

Despite the quality of the images, the JSEG algorithm is efficient in omnidirectional images but the repetition of *pixel*s in the panoramic scenes can lightly change the results compared with the same original image. It is necessary to implement a previous step for rectification.

Although, the combination of cameras and mirrors are slightly used in agricultural environments, it is possible to evolve the image acquisition step. It is a fact that the catadioptric systems have immense potential due to the amount of data from the environment, restricted in its conception. Acquiring images with good spatial/graphical resolution of the environment may work with any known algorithms applied in navigation tasks through the rectification process.

6. FUTURE WORKS

The next step is to maximize the process of image acquisition and change the physical support inside structure for the camera. Therewith, it is intend to get more good results with JSEG algorithm, in which optimized parameters values for the algorithm turn reliable.

A Naive Bayes probabilistic classifier will be implemented, which is responsible for classifying the mapped homogeneous regions in JSEG. *A priori*, the classifier training process, required for color quantization step, will after be used as key-features.

7. ACKNOWLEDGEMENTS

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

8. REFERENCES

Deng, Y., Manjunath, B. S. and Shin, H., 1999. “Color Image Segmentation”. Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, CVPR’99, Fort Collins, CO, vol.2.

Ericson, S. and Astrand, B., 2010. “Row-detection on an agricultural field using omnidirectional camera,” IEEE/RSJ International Conference on Intelligent Robots and Systems.

Fiala, M. and Basu, A., 2002. “Robot navigation using panoramic landmark tracking”. IEICE Transactions on Information and Systems.

Grassi Jr, V. and Okamoto Jr., J., 2007. “Development of an omnidirectional vision system”. Mechatronics and Machine Vision in Practice.

Li, M., Imou, K., Wakabayashi, K. and Yokoyama, S., 2009. “Review of research on agricultural vehicle autonomous guidance”. International Journal of Agricultural and Biological Engineering - IJABE.

Li, M., Liu, Z. H., Huang, J., Dai, S. H. and Weng, W., 2011. “Position estimation method for agricultural field road”. International Conference on Agricultural Engineering (ICAE).

Lulio, L. C., Tronco, M. L. and Porto, A. J. V., 2011. "Omnidirectional vision system for agricultural mobile robot navigation". Congresso Nacional de Engenharia Mecânica - CONEM/ABCM. August 18 - 21, 2010, Campina Grande, Paraíba, Brazil.

Lulio, L. C. 2011. "Técnicas de visão computacional aplicadas ao reconhecimento de cenas naturais e locomoção autônoma em robôs agrícolas móveis". 353p. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, SP, Brazil.

Nayar, S. K. and Baker, S., 1997. “Catadioptric image formation”, DARPA Image Understanding Workshop.

Scaramuzza, D., 2008. Ominidirectional vision from calibration to robot motion estimation. Tese de Doutorado. Universita di Perugia.

Svoboda, T., Pajdla, T. and Hlavac, V., 1997. “Central panoramic cameras: Geometry and design”. Praga: Czech Technical University, Faculty of Electrical Engineering, Center for Machine Perception.

Trentini, V. B., Rodrigues, V. A., Marana, A. N., Lulio, L. C. and Tronco, M. L., 2013. "Stereo Matching of Natural Scenes on External Agricultural Environment", IX Workshop de Visão Computacional.

Yuanjie, Z., Jie, Y., Yue, Z. and Yuzhong, W., 2006 "Color-texture base unsupervised segmentation using JSEG with fuzzy connectedness". In Journal of Systems Engineering and Electronics, vol. 17, pp. 213-219.

Zhu, Z., Rajasekar, K., Riseman, E. and Hanson, A., 2000. “Panoramic virtual stereo vision of cooperative mobile robots for localizing 3d moving objects”. In IEEE Workshop on Omnidirectional Vision, pp. 29–36.

9. RESPONSIBILITY NOTICE

The author(s) is (are) the only responsible for the printed material included in this paper.