Portable 3D laser-camera calibration system with color fusion for SLAM

Javier Navarrete-Sanchez
Instituto Universitario de Investigación en Informática, University of Alicante
Alicante, Spain
e-mail: javi.navarrete@ua.es

Diego Viejo
Instituto Universitario de Investigación en Informática, University of Alicante
Alicante, Spain
e-mail: dviejo@dccia.ua.es

Miguel Cazorla
Instituto Universitario de Investigación en Informática, University of Alicante
Alicante, Spain
e-mail: miguel.cazorla@ua.es

***Abstract*—Nowadays, the use of RGB-D sensors have focused a lot of research in computer vision and robotics. These kinds of sensors, like Kinect, allow to obtain 3D data together with color information. However, their working range is limited to less than 10 meters, making them useless in some robotics applications, like outdoor mapping. In these environments, 3D lasers, working in ranges of 20-80 meters, are better. But 3D lasers do not usually provide color information. A simple 2D camera can be used to provide color information to the point cloud, but a calibration process between camera and laser must be done. In this paper we present a portable calibration system to calibrate any traditional camera with a 3D laser in order to assign color information to the 3D points obtained. Thus, we can use laser precision and simultaneously make use of color information. Unlike other techniques that make use of a three-dimensional body of known dimensions in the calibration process, this system is highly portable because it makes use of small catadioptrics that can be placed in a simple manner in the environment.**

We use our calibration system in a 3D mapping system, including Simultaneous Location and Mapping (SLAM), in order to get a 3D colored map which can be used in different tasks. We show that an additional problem arises: 2D cameras information is different when lighting conditions change. So when we merge 3D point clouds from two different views, several points in a given neighborhood could have different color information. A new method for color fusion is presented, obtaining correct colored maps. The system will be tested by applying it to 3D reconstruction..

Keywords:- 2D-3D calibration; RGB-D information; color fusion; SLAM.

I. INTRODUCTION

2D information currently collected by video and conventional cameras is insufficient for a user familiarized to visualize 3D data. Usually, video games started the three-dimensional representation of virtual worlds. In robotics, recently there are a lot of applications using 3D data. Realistic 3D models obtained from real environments and structures are becoming research line with multiple possibilities.

There are multiple ways to obtain 3D information from the environment. For example, time-of-flight or stereo cameras allow getting that information, but they have a very limited range and a high associated noise. Stereo cameras also suffer from the lack of texture. 3D range lasers are a more accurate kind of sensors. The range of this sensors can vary from a few meters to hundreds of meters, with errors of the order of millimeters to centimeters. Despite their geometric precision, the use of laser does not provide color information of the scene and this feature represents a deficiency with respect to other devices. However, certain lasers can provide us additional information about the intensity of reflection of light on certain materials. This information depends on the material from which the object is made. A glass, depending on the angle of incidence, does not offer any intensity value while a mirror would give a maximum value.

In this paper, we use the most accurate information provided by a laser, in conjunction with the color information of a regular digital camera, in order to obtain a colored 3D point cloud. We present here a procedure to perform the calibration between 3D laser data and a 2D camera image. This procedure finds extrinsic parameters for the camera-laser system by using laser intensity information. The most brilliant points in 3D data set are matched to the corresponding points in the image. The calibration process is performed using this information. We automatically find the transformation between the camera and the laser. Once the calibration process is done, the pixel color mapping to 3D point is done automatically. The proposed system, compared to existing ones, is completely portable, can be performed anywhere and does not change the environment. The validity of the system will be conducted using a 3D matching method which will reconstruct the 3D environment around the robot.

The rest of the paper is structured as follows: first we review the state of the art and describe the hardware devices used. We continue with the description of the method used for calibration. Then, the method used to calculate the 3D registration between different poses is presented; and finally some conclusions are drawn.

II. STATE OF THE ART

Currently, due to current power of computers and the use of graphics cards, robotics has begun to use 3D data, both in mapping and its use in solving the problem of SLAM (Simultaneous Location And Mapping). The use of color in the 3D points can help in these processes.

At present there are different ways for obtaining 3D data with color information. On the one hand we find works dealing with the problem from the use of images. The use of time of flight or stereo cameras provides color to the 3D points. However, these cameras have limited range and are highly noisy.

A second way is the combined use of 3D laser devices and cameras. In this case, a calibration is necessary to find correspondences between the points of the camera and the 3D device. In the works [2] and [6] a three dimensional object with known geometry is used, for a previous calibration, from which control points are obtained, which are used to find correspondences. This object is large and hardly portable.

On the other hand, work [4] uses a laser device that offers great precision in capturing the intensity of reflection of laser materials. Using this feature the authors generate a point cloud colored in gray scale with high accuracy, because the color assignment to each point is immediate and the color assignment is invariant to changes in light, being able to even perform it in complete darkness.

In our method, we use the intensity readings in order to obtain a calibration process in the correspondence between the laser and image, and moreover the actual color of the camera for coloring the point cloud. So in our case, we need not know the geometry of any object for calibration. The elements used in our case are easily adaptable to any environment and do not require a great placement, making the system highly portable.

The composition and fusion of multiple image data is a problem that has been widely investigated and lots of proposals had been made. Brown and Lowe [23] proposed an automatic complete process to create a panoramic view from multiple images using different techniques of alignment and image fusion (blending). Burt and Adelson [24] defined the methods *for multiband blending*, used in several algorithms for image fusion. Their proposal was a multirresolution spline technique for combining two or more images into a larger image mosaic. In their procedure, the images to be splined are first decomposed into a set of band-pass filtered component images.

In our proposal, we adapt this technique to be used with 3D data, using camera captures as data inputs, and obtaining a color-combined 3D point cloud.

For the experiments, we use Point Cloud Library (PCL) [25] as it is frequently updated and is widely used in 3D data processing research.

III. Experimental Setup

As mentioned above, the calibration process allows the matching of each of the readings from a laser with the pixels of an image. This is done by using the devices mentioned below.

Fig. 1. Hardware system used.

- Laser sensor SICK LMS-200.
- Swinging unit PowerCube.
- Digital cameral Pentax Optio SR.

As shown in Fig. 1, the laser device is mounted on a swinging unit PowerCube, which is mounted on a PowerBoot mobile robot. Both laser and camera are fixed in the robotic swing arm so that both elements move simultaneously. The mobility of the swing arm allows us to get readings at different angles, obtaining a complete 3D point cloud of the environment. The image from the camera is captured when the swinging arm is at the initial tilt position.

IV. 3D Camera-Laser Calibration

The procedure for calibration is as follows:

- Data collection. Obtaining the 3D point cloud, with certain points easily identifiable by their intensity. Take a snapshot from the 2D camera.
- Transformation to a 2D image projection from the 3D point cloud.
- Segmentation of the control points in both 2D images.
- Estimation of the transformation between two images. A correspondence must be made between the points detected.
- Apply correspondence between image pixels and 3D points to assign color to the 3D data.

A. Data capture

In this step, just do the data collection of data both 3D points and 2D image. It is necessary to establish some correspondence between points in both sets of data. To do this, we use a reflective material that allows to be captured by the laser with high intensity values, while in the 2D image is also easily readable. We use a reflective catadioptric to set these points. So we get a cloud with certain points whose intensity values are far from the rest of the environment.

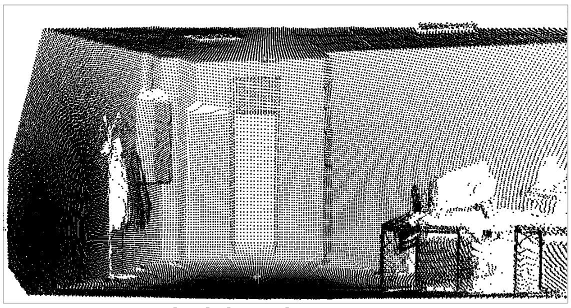


Fig. 2. 3D point cloud obtained.

Once the 3D point cloud from the swinging is taken, we take a picture from the camera. Obviously the field of view of the current camera attached to the swinging arm is much smaller than the field of laser, so in the calibration process we ensure that the position of the catadioptrics are captured by the laser and then camera.

Fig. 3. 3D point cloud obtained.

B. Obtaining a 2D Image from a 3D Point Cloud

To generate a 2D image from a 3D point cloud we use a *Pinhole* camera model:

$$u_{3D} = f * (X/Z)$$
$$v_{3D} = f * (Y/Z)$$

where (X, Y, Z) are the coordinates of 3D point, (u_{3D}, v_{3D}) are the coordinates of the point in the 2D image generated and f the focal distance. Using this camera model, a 2D image is constructed using as pixel color the values obtained by the laser intensity (normalized to grayscale): Sick laser models provide a reflectance intensity value associated with the distance in a given beam. Taking into account that in our system, most of the reading materials provide 0 values except reflective materials, the result is a black image, except for the control points (the catadioptrics). SICK LMS 200 laser characteristics provide only 4 intensity values, and although there are lasers on the market with more intensity values, this is enough to detect the control points for the proposed technique.

C. Segmenting the Control Points in both Images

The use of catadioptric gives us the advantage that these are clearly identifiable. Its reflectivity exceeds that of the vast majority of the surrounding materials, which facilitates their identification in the intensity image generated from the laser readings. Besides, due to the characteristics of reflective materials, they usually have a color very significant in an image (red, orange, yellow), not commonly found in the environment. At indoor environments, the use of the camera flash makes simpler to identify them.

On the other hand, the use of these elements makes this calibration method a very simple and portable and can be installed quickly and anywhere, without having to modify the environment.

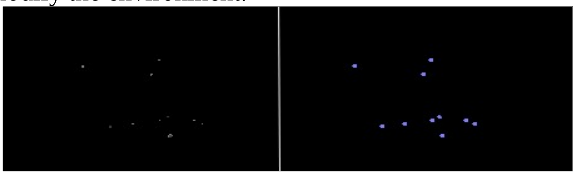


Fig. 4. Section 2D image extracted from the cloud of points. Left: intensity image. Right: identification of the reflectors.

The segmentation procedure is the same for both images, although only some parameters are changed for segmentation. The segmentation process is as follow [13]:

Segmentation of HSV values. First, we identify the values H (hue) and S (saturation) for catadioptrics and then segmentation is performed. In the case of the laser, is much simpler, since all points return a zero value, except the points reflected by the catadioptric.

Application of morphological operators (opening, closure) for consistent areas.

Application of region detection algorithm (*Blobs*), forcing these to have a minimum size. This will avoid the appearance of two control points on the same catadioptric. One the blob is detected, its centroid is calculated.

Matching checkpoints. In our case, deformation does not occur; a match can be made directly using the coordinates of the catadioptric in the image. In the case of omnidirectional images from cameras, fisheye or others, it would require a more efficient matching, like using graphs.

D. Calculating the Transformation between both Images

To obtain the transformation performed between the two images, we obtain the transformation matrix T between the control points. It is not an affine transformation, but a projective. A projective transformation is a combination of a rotation, translation, scaling and nonlinear component. The affine transformation is a particular case of the projective.

Let (x', y') be the points from a plane and (x, y) the one from the other plane, one can get the relationship between the two points by the equation:

$$\begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} = T * \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

where

$$x' = \frac{x'_1}{x'_3}, \quad y' = \frac{x'_2}{x'_3}, x = \frac{x_1}{x_3}, y = \frac{x_2}{x_3}$$

For a projective transformation, the transformation matrix is defined as:

$$T = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ v_1 & v_2 & k \end{bmatrix}$$

where v_1 and v_2 are the components of the nonlinear transformation and k is dependent on other terms. So we need to calculate the coefficients c_{ij} and v_i. To do this, we need at least 4 pairs of matched points thus 8 equations with 8 unknowns must be solved in most cases, using least squares [11]. We assume homogeneous coordinates (x, y) with $x_3 = 1$, so we have:

$$x' = \frac{c_{11}x + c_{12}y + c_{13}}{v_1 x + v_2 y + k}$$

$$y' = \frac{c_{21}x + c_{22}y + c_{23}}{v_1 x + v_2 y + k}$$

Under that approach more correspondences can be used that to approximate the solution, but this does not ensure that the approximation found is correct for all points. In some special cases, if we take 4 points and 3 pairs of them are in the same line, finding the solution is not possible. This fact must be taken into account in the placement of the catadioptric.

After calculating T each pixel from the camera can be corresponded to one pixel from the 2D image obtained from 3D point cloud. This picture represents all the readings done with the laser 3D. As 3D laser and the 2D camera have not equal field of view, some 3D points will be not colored.

V. COLOR FUSION

The obtained point cloud has some color irregularities caused by light changing or *vignetting* (less intensity at borders image). This effect is specially intensive in areas of the cloud where the color must be uniform (as a wall or floor).

In order to mitigate this effect we propose a method based in the well know blending technique for 2D images[24].

We define each point P as:

$P=(x,y,z,r,g,b,w)$

where x,y,z are the point coordinates, r,g,b are the color components, and w the weight assigned to this point.

Each point assigns his weight in function the distance to the center of the image 2D. In each captured image, we give the max value 1 to pixels in the center of the image, and 0 in borders, using a progressive linear function for the rest.

The assignment give more saliency to the center of the images, but it is possible to modify this criterion; assuming for example a specific region of the image.

The weight assignation method is given to the pixels in the 2D image, and when the pixel is mapped to a 3D point (as described before in the calibration method) this point saves the correspondent weight.

After reconstruction, all points are evaluated using the neighbors inside a specified radius r. Of course the result for each point evaluation is stored in a new point cloud to avoid overlapping in the color calculation.

Each point calculates each color band value (I) as follows:

$$I_p = \frac{\sum_{\Delta i \in r} I_i w_i}{\sum_{\Delta i \in r} w_i}$$

The final point color is the combination of the obtained band values. The radius parameter controls the number of points which are used in.

We have evaluated the possibility of managing the radius parameter dynamically, using a greater value in areas of the cloud where the point density is low, and a smaller value in areas with high density. The initial radius is a small value, and then if the number of points in the set is lower than a limit, the radius is increased. This ensures a minimum number of elements inside the set of study.

Note that this process can be used to coloring voxels, if we use the points inside a voxel instead of a specified radio. In this case, the color of the resultant voxel is a mixture of all colors inside the voxel.

VI. RESULTS

Some results can be visualized in the two following figures. Obviously those are 3D points that are within the field of view of the camera are colored.

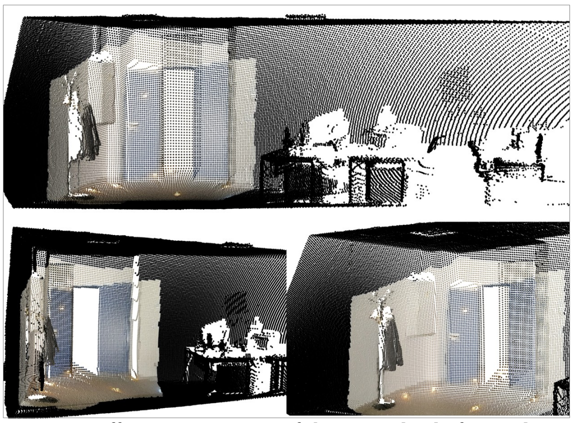


Fig. 5. Different perspectives of the point cloud after application of our method.

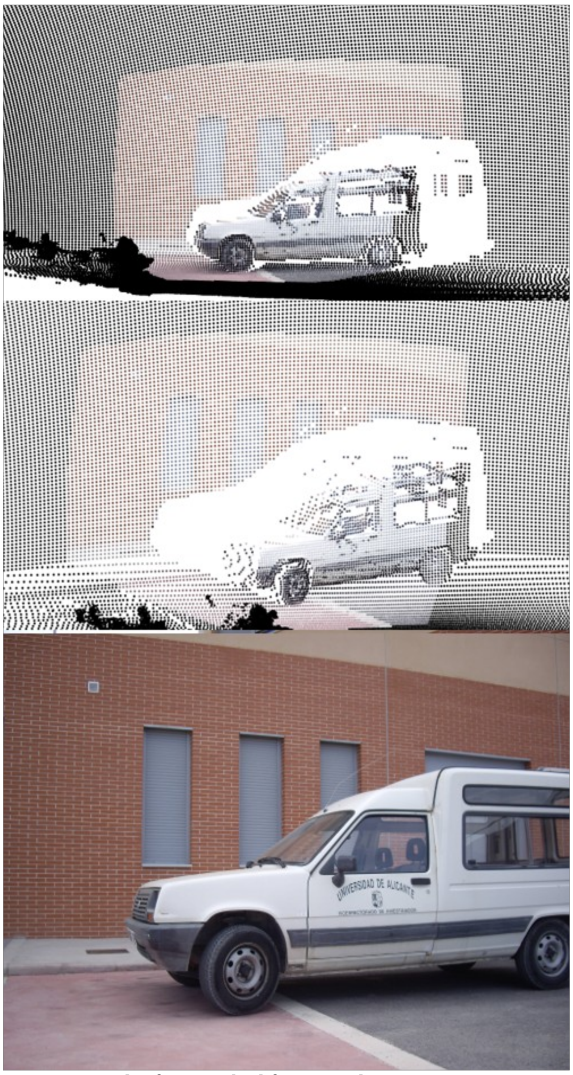
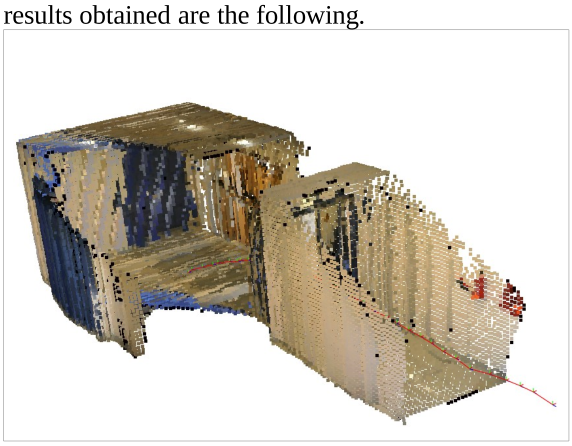


Fig. 6. Result of our method for an outdoor scene.

In order to get a 3D reconstruction (map) of the sequence, we use the method proposed in [22]. The results obtained are the following.



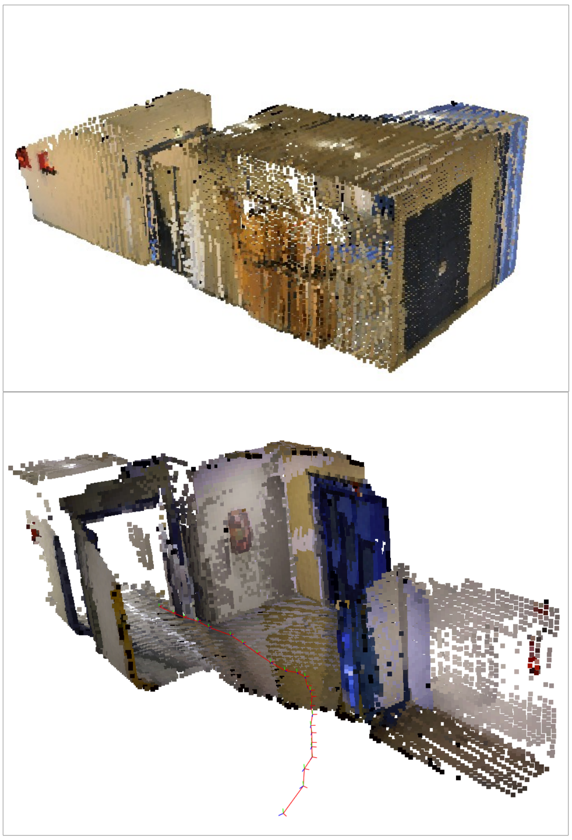


Fig. 7. Reconstruction obtained: the first two images correspond to a sequence and the bottom to a different one. The red lines indicate the robot path.

In the Fig.7 we can observe the effects of light changing. Some color bands have appeared in the walls and they do not have a uniform color. We applied our process to the same point cloud and we appreciate how these effects disappear significantly in Fig.8

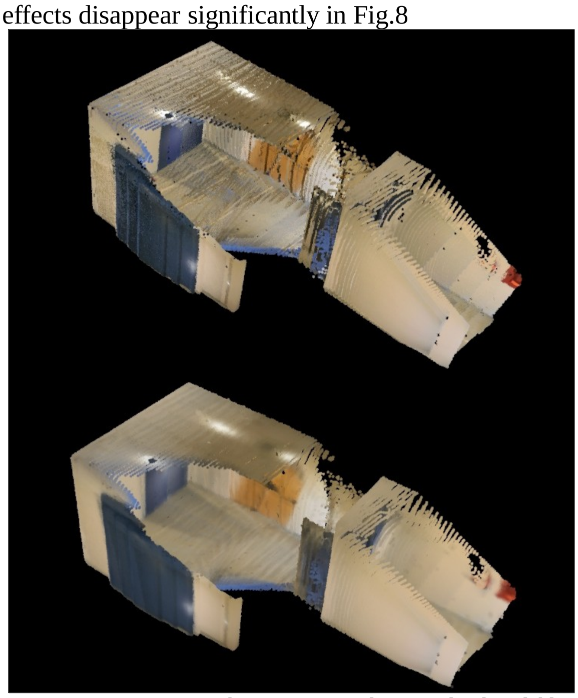


Fig. 8. Comparison between original point cloud and blended point cloud.

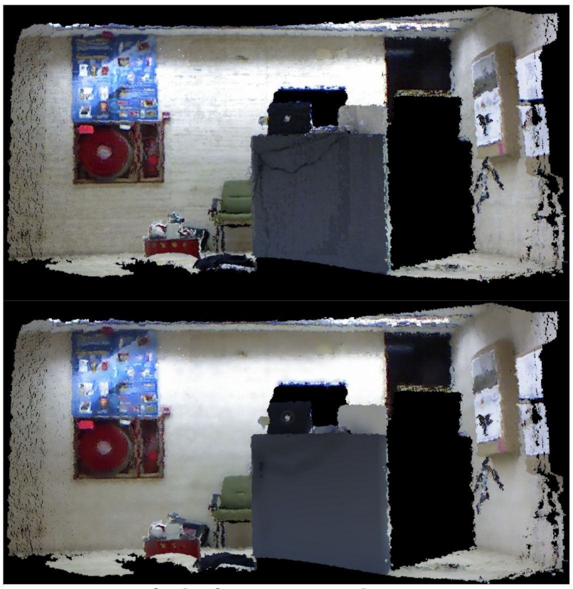


Fig. 8. Use of color fusion in Kinect datasets.

Optimal radius, depend on point cloud density and light changing effects.

VII. Conclusions and Future Works

We have presented a portable 3D laser and 2D camera calibration system. The system uses catadrioptics in order to match points from the 3D point cloud and the 2D image. The use of catadrioptics is useful to automatically match points, using computer vision techniques. The result of the method has been tested using a 3D registration method, obtaining good result in 3D mapping.

It seems that the resulting accuracy is quite accurate, although not defined quantitative method is defined to measure the error. For this reason, we plan to define a ground truth system in order to test this error. The use of a fisheye or omnidirectional camera will provide more points, but a deeper study must be done and the use of a graph structure must be used in order to match 3D and 2D points.

Our color data fusion reduces light changing effects, in photorealistic reconstruction.

In order to improve the coloring results from multiple data images and reduce blur effects, we are planning to use color histograms and to study different methods for combining this information with color fusion.

Acknowledgments

This work has been supported by project DPI2009-07144 from Ministerio de Investigacion, Ciencia e Innovacion (Spain) and GRE10-35 from Universidad de Alicante (Spain).

References

[1] M. Sonka, V. Hlavac y Roger Boyle, Image Processing, Analysis, and Machine Vision, Second Edition. Editorial PWS publishing, 1999.

[2] Kai Pervz, Andreas Nchter, Hartmut Surmann and JoachimHertzberg, Automatic Reconstruction of Colored 3D Models. In proceedings of robotik 2004 (p. 215-222).

[3] Karsten Scheibe1, Martin Scheele1, and Reinhard Klette. Data Fusion and Visualization of Panoramic Images and lser Scans

[4] P.Dias, V. Sequeira, J.G.M. Gonalves, F.Vaz. Automatic registration of lser reflectance and colour intensity images for 3D reconstruction. Robotics and Autonomous Systems 39 (2002) 157-168.

[5] P. Allen and I. Stamos. Integration of range and image sensing forphotorealistic 3d modelling. In Proc of IEEE Inter. Conf. on Robotics and Automation, 2000.

[6] Derik Schroeter, Paul Newman. On the Cross-Calibration of a Camera and a 3D-lser Range Finder. 2nd SEAS DTC Technical Conference - Edinburgh 2007.

[7] A low cost 3D scanner based on structured light C. Rocchini, P. Cignoni, C. Montani, P. Pingi and R. Scopignoy Computer Graphics Forum, Volume 20, Issue 3, pages 299308, September 2001

[8] Viejo, D. and Cazorla, M. 3D Feature Extraction and Modelling for SLAM. 39th International symposium on robotics (IRS08). Corea. 2008

[9] Richard Hartley and Andrew Zisserman ,Multiple View Geometry in Computer Vision. Candbridge University Press 2000, 2003. ISBN: 0521623049.

[10] Oficial Matlab web page. http://www.mathworks.com

[11] Goshtasby Ardeshir, Image registration by local approximation methods, Image and Vision Computing, Vol. 6, 1988, pp. 255-261.

[12] P. Besl, N. McKay, A method for registration of 3-d shapes, IEEE Trans. on Pattern Analysis and Machine Intelligence 14 (2) (1992)

[13] E. Trucco y A. Verri, Introductory Techniques for 3-D Computer Vision Prentice Hall, 1998

[14] R. Klette, K. Schlns y A. Koschan, Computer Vision, Three-Dimensional Data from Images. Editorial Springer, 2001.

[15] Yi Ma, Stefano Soatto, Jana Kosecka, S.Shankar Sastry. An Invitation to 3-D Vision. Springer, 2005.

[16] D. L. Page, Y. Sun, A. F. Koschan, J. Paik, M. A. Abidi, Normal vector voting: crease detection and curvature estimation on large, noisy meshes, Graph. Models 64 (3/4) (2002).

[17] D. M. Cole, A. R. Harrison, P. M. Newman, Using naturally salient regions for slam with 3d laser data, in: Proc. of the IEEE ICRA, 2005.

[18] N. J. Mitra, A. Nguyen, Estimating surface normals in noisy point cloud data, in: SCG 03: Proceedings of the nineteenth annual symposium on Computational geometry, ACM Press, 2003, pp..

[19] M. Martn, J. Gmez, E. Zalama, Obtaining 3d models of indoor environments with a mobile robot by estimating local surface directions, Robotics and Autonomous Systems 48 (2-3) (2004).

[20] F.Y.Shih, S.Cheng, Automatic seeded region growing for color image segmentation, Image and Vision Computing 23 (2005).

[21] J. Fan, M. Zeng, M. Body, M. S. Hacid, Seeded region growing: an extensive and comparative study, Pattern Recognition Letters 26 (2005)

[22] Diego Viejo, Miguel Cazorla. 3D plane-based egomotion for SLAM on semi-structured environment . 2007 IEEE International Conference on Intelligent Robots and Systems. San Diego, California. Estados Unidos. 2007

[23] Peter J. Burt. Edward H. Adelson. A Multiresolution Spline With Application to Image Mosaics. ACM Transactions on Graphics (TOG) Vol 2. 1983

[24] Matthew Brown and David G. Lowe. Automatic Panoramic Image Stitching using Invariant Features. International Journal of Computer Vision, 2007

[25] http://pointclouds.org/ Web page of the Point Cloud Library project. Last acess 04/04/2012