Portable 3D Laser-Camera Calibration System with Color Fusion for SLAM

Javier Navarrete-Sanchez, Diego Viejo* and Miguel Cazorla

Instituto Universitario de Investigación en Informática, University of Alicante Alicante, Spain

Abstract: Recently, the use of RGB-D sensors has been the focus of a lot of research in computer vision and robotics. These kinds of sensors, like Kinect, are able 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. However, 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. As such, we can use laser precision and simultaneously make use of color information. Unlike other techniques that make use of a 3D bodies 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 camera 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 is tested by applying it to 3D reconstruction.

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

Introduction

2D information currently collected by video and conventional cameras is insufficient for a user familiarized to visualize 3D data. Historically, video games started the 3D representation of virtual worlds. In robotics, recently there have been many applications using 3D data. Realistic 3D models obtained from real environments and structures are becoming a 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 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 these 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 lasers does not provide color information of the scene, which represents a deficiency compared 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 piece of 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 a 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 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.

State Of The Art

Currently, owing to the 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. There exist, however, works dealing with the problem from the use of images [1]. The use of time of flight or stereo cameras provides color to the 3D points. However, these cameras have limited range and are very noisy. Another method used is the combined use of 3D laser devices and cameras [2]. In this case, a calibration is necessary to find correspondences between the points of the camera and the 3D device. In [3] and [4], a three dimensional object with known geometry for a previous calibration is used to obtain control points, which are used to find correspondences. This object is large and hardly portable.

Another approach [5] 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 is invariant to changes in light. It is even able to be performed in complete darkness.

In our method, we use intensity readings in order to obtain a calibration process in the correspondence between the laser and image, as well as the actual color from 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 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 [6] 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 [7] defined the methods for multiband blending used in several algorithms for image fusion. Their proposal was a multiresolution 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 following experiments, we used Point Cloud Library (PCL) [8] as it is frequently updated and is widely used in 3D data processing research.

Diego Viejo obtained his BSc and MSc in Computer Science in 2002 and his PhD in 2008 both from the University of Alicante. Since 2004, he has been a lecturer and a researcher in the Department of Computer Science and Artificial Intelligence (DCCIA) at the University of Alicante. His research interests are focused on 3D vision applied to mobile robotics.

Miguel Cazorla received a BS degree in Computer Science from the University of Alicante (Spain) in 1995, and a PhD in Computer Science from the same university in 2000. He is currently an associate professor in the Department of Computer Science and Artificial Intelligence at the University of Alicante. He has done several postdocs stays: ACFR at University of Sydney with Eduardo Nebot, IPAB at University of Edinburgh with Robert Fisher, CMU with Sebastian Thrun and SKERI with Alan Yuille. He has published several papers on robotics and computer vision. His research interests are focused on computer vision and mobile robotics (mainly using vision to implement robotics tasks).

Javier Navarrete obtained his BSc in Computer Science in 2002 and his Master's in 2012, both from the University of Alicante. Since 2008, he has been a technician and a researcher in the University Institute for Computing Research (IUCR) at the University of Alicante. His research interests are focused on 3D vision and robotics.

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 listed below:

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

As shown in Figure 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.

Figure 1. Hardware system used.

3D Camera-Laser Calibration

The procedure for calibration is as follows:

· Obtaining the 3D point cloud with certain points easily identifiable by their intensity. Take a snapshot with 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.
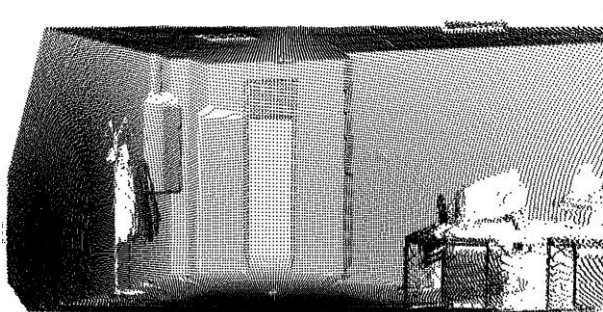· Apply correspondence between image pixels and 3D points to assign color to the 3D data.

Figure 2. 3D point cloud obtained.

Data capture

This step involves collecting the data of both the

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 is able to be captured by the laser with high intensity values, while the 2D image is also easily readable. We use a reflective catadioptric system to set these points. Doing so, we get a cloud with certain points with intensity values far from the rest of the environment.Once the 3D point cloud from the swinging PowerCube unit is taken, we take a picture with the camera. Obviously, the field of view of our camera attached to the swinging arm is much smaller than the field of a laser, so in the calibration process we ensure that the position of the catadioptrics are captured by the laser and then the camera.

Figure 3. 3D point cloud obtained.

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:

$$u3D = f \times (X / Z)$$
$$v3D = f \times (Y / Z),$$

where (X, Y, Z) are the coordinates of the 3D point, $(u3D, v3D)$ are the coordinates of the point in the 2D image generated, and f is 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 four 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.

Segmenting the Control Points in both Images

The use of catadioptrics gives us the advantage that the control points in both images are clearly identifiable. The 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 very identifiable color in an image (red, orange, yellow) not commonly found in the environment. In indoor environments, the use of the camera flash makes identifying them simpler. Moreover, the use of these elements makes this calibration method very simple and portable, and the system can be installed anywhere quickly, without having to modify the environment.
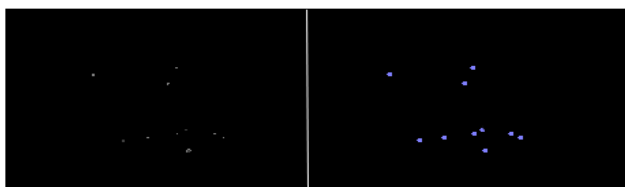

Figure 4. Section 2D image extracted from the point cloud. Left: intensity image. Right: identification of the reflectors.

The segmentation procedure is the same for both images, although only some parameters are changed for segmentation. The segmentation process is as follows [9] :

- Segmentation of HSV values. First, we identify the values H (hue) and S (saturation) for the catadioptrics, and then segmentation is performed. The case of the laser is much simpler since all points return a zero value, except the points reflected by the catadioptrics.
- Application of morphological operators (opening, closure) for consistent areas.
- Application of region detection algorithm (Blobs), forcing them to have a minimum size. This will avoid the appearance of two control points on the same catadioptric image. Once 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 catadioptrics in the image. In the case of omnidirectional images from cameras, fisheye or others, it would require a more efficient matching, like using graphs.

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.

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

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

where

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

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

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

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 four pairs of matched points, giving eight equations with, in most cases, eight unknowns that must be solved. This is done using least squares [10]. We assume homogeneous coordinates (x,y) with $x_3 = 1$, giving us:

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

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

With this approach, more correspondences can be used 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 four points and three 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 catadioptrics.

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

Color Fusion

The obtained point cloud has some color irregularities caused by light changing or vignetting (less intensity at image borders). This effect is especially 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-known blending technique for 2D images [7].

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 is the weight assigned to this point.

Each point assigns its weight in function as the distance to the center of the 2D image. In each captured image, we give the max value of 1 to pixels in the center of the image, and zero in borders. A progressive linear function is then used for the rest. This assignment gives 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 assigning method is given to the pixels in the 2D image, and when a 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:

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

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

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 studied set.

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

Results

Results can be visualized in the two following figures. The 3D points that are within the field of view of the camera are colored.
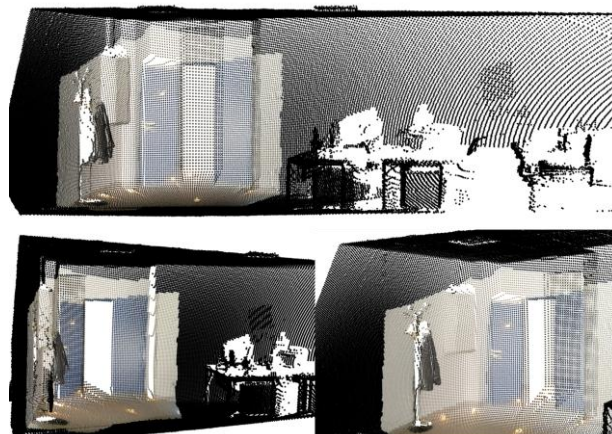

Figure 5. Different perspectives of the point cloud after application of our method.
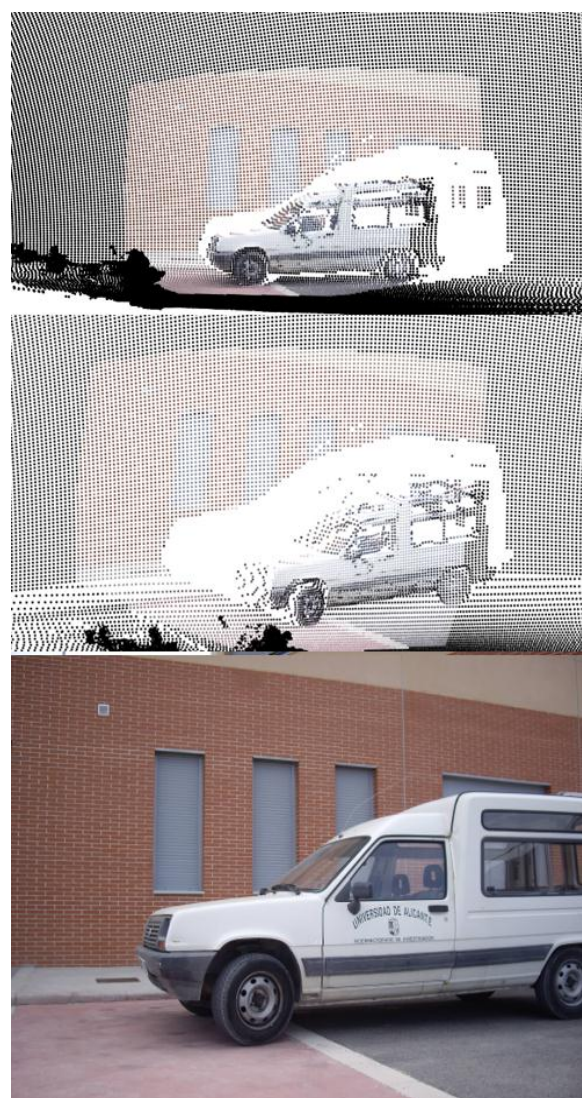

Figure 6. Result of our method for an outdoor scene.

In order to get a 3D reconstruction (map) of the sequence, we use the method proposed in [11], which uses the feature extraction method from [12]. The results obtained are shown below.

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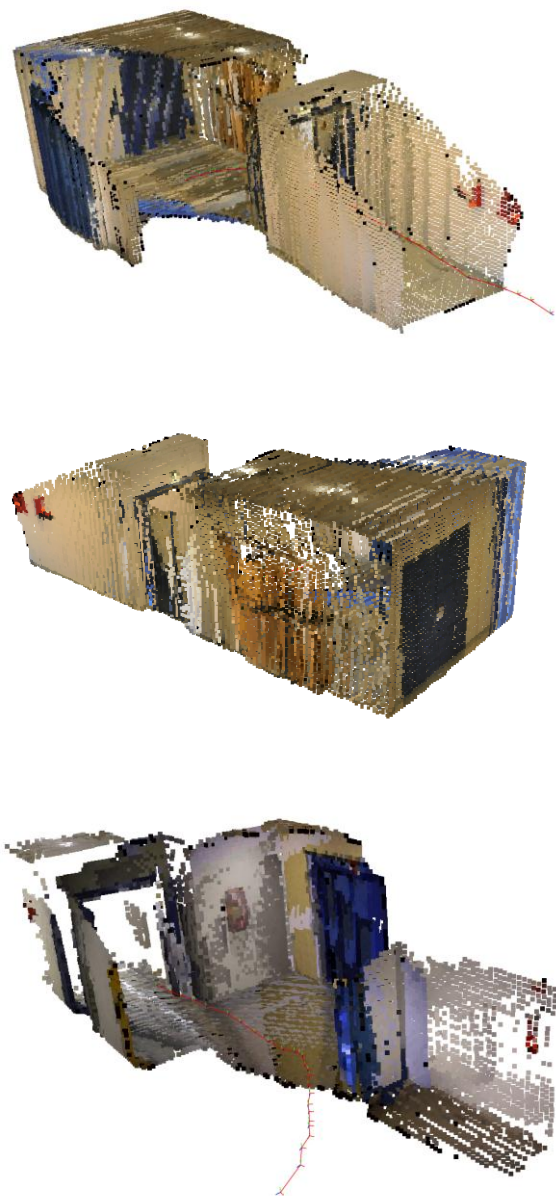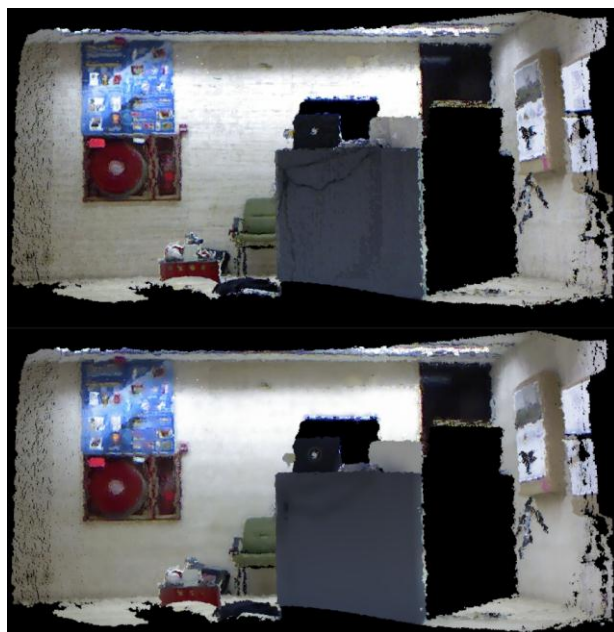

Figure 9. Use of color fusion in Kinect datasets.

Figure 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 Figure 7, we can observe the effects of changing light. 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 these effects disappear significantly (Figure 8).

The optimal radius for our method depends on point cloud density and light changing effects.
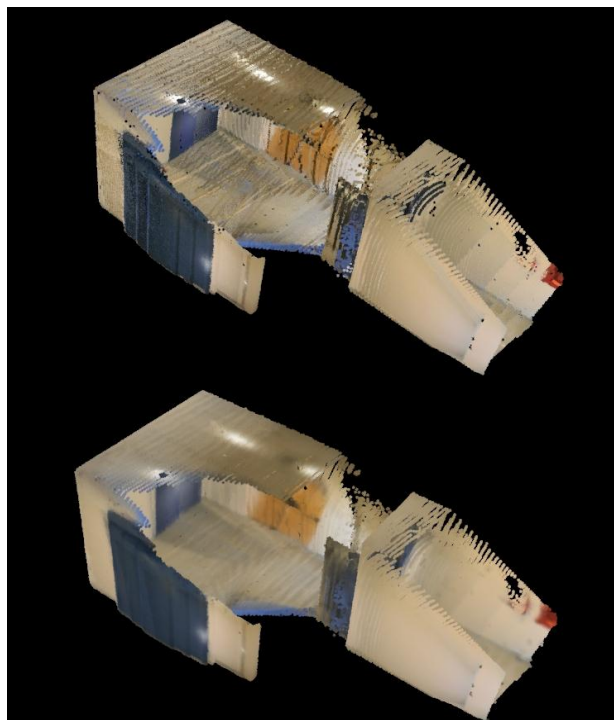
Conclusion and Future Work

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

mapping. It seems that the results are quite accurate, although a defined quantitative method is not used to measure the error. For this reason, we plan to define a ground truth system in order to test for this error. Also, the use of a fisheye or omnidirectional camera would provide more points, but an in-depth study must be done and 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

Acknowledgement

References

[1] K. Scheibe, M. Scheele, and R. Klette, "Data fusion and visualization of panoramic images and laser scans," in *ISPRS working group V/1, Panoramic Photogrammetry Workshop*, Dresden, Germany, 2004, vol. XXXIV, part 5/W16. Available: http://www.isprs.org/proceedings/XXXIV/5-W16/papers/PanoWS_Dresden2004_Scheibe.pdf

[2] I. Stamos and P. K. Allen, "Integration of range and image sensing for photo-realistic 3d modeling," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2000, vol. 2, pp. 1435-1440
doi: 10.1109/ROBOT.2000.844799

[3] K. Pervölz, A. Nüchter, H. Surmann, and J. Hertzberg, "Automatic reconstruction of colored 3d models," in *Robotik*, Düsseldorf, Germany, 2004, pp. 215-222.

[4] D. Schroeter and P. Newman, "On the cross-calibration of a camera and a 3d-laser range finder," in *2nd SEAS DTC Technical Conference*, Edinburgh, Scotland, 2007.

[5] P. Dias, V. Sequeira, J. G. M. Gonalves, and F.Vaz, "Automatic registration of laser reflectance and colour intensity images for 3d reconstruction," *Robotics and Autonomous Systems,* vol. 39, no. 3-4, pp. 157-168, 2002.
doi: 10.1016/S0921-8890(02)00201-4

[6] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Transactions on Graphics,* vol. 2, no. 4, pp. 217-236, 1983.
doi: 10.1145/245.247

[7] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *International Journal of Computer Vision,* vol. 74, no. 1, pp. 59-73, 2007.
doi: 10.1007/s11263-006-0002-3

[8] *Web page of the point cloud library project*, [Online].
Available:
http://pointclouds.org/ [Accessed: April, 2012]

[9] E. Trucco and A. Verri, *Introductory techniques for 3-d computer vision*. Upper Saddle River, NJ: Prentice Hall, 1998.

[10] A. Goshtasby, "Image registration by local approximation methods," *Image and Vision Computing,* vol. 6, no. 4, pp. 255-261, 1988.
doi: 10.1016/0262-8856(88)90016-9

[11] D. Viejo and M. Cazorla, "3d plane-based egomotion for slam on semi-structured environment," in *IEEE International Conference on Intelligent RObots and Systems (IROS)*, San Diego, California, 2007, pp. 2761-2766.
doi: 10.1109/IROS.2007.4399138

[12] D. Viejo and M. Cazorla, "3d feature extraction and modelling for slam," in *39th International symposium on robotics (IRS08)*, Korea, 2008.