LIDAR and Panoramic Camera Extrinsic Calibration Approach Using a Pattern Plane

Angel-Iván García-Moreno, José-Joel Gonzalez-Barbosa, Francisco-Javier Ornelas-Rodriguez, Juan B. Hurtado-Ramos, and Marco-Neri Primo-Fuentes

CICATA, Instituto Politécnico Nacional
Cerro Blanco 141 Col. Colinas del Cimatario, Quéretaro
jgonzalezba@ipn.mx

Abstract. Mobile platforms typically combine several data acquisition systems such as lasers, cameras and inertial systems. However the geometrical combination of the different sensors requires their calibration, at least, through the definition of the extrinsic parameters, i.e., the transformation matrices that register all sensors in the same coordinate system. Our system generate an accurate association between platform sensors and the estimated parameters including rotation, translation, focal length, world and sensors reference frame. The extrinsic camera parameters are computed by Zhang's method using a pattern composed of white rhombus and rhombus holes, and the LIDAR with the results of previous work. Points acquired by the LIDAR are projected into images acquired by the Ladybug cameras. A new calibration pattern, visible to both sensors is used. Correspondence is obtained between each laser point and its position in the image, the texture and color of each point of LIDAR can be know.

Keywords: panoramic camera, LIDAR, sensor calibration, extrinsic calibration.

1 Introduction

Reconstruction of urban environments is a challenge given the variety of scenes that can be scanned and problems that arise when working with real data. Many techniques are used for three-dimensional reconstruction, each with their own limitations and none of them efficiently solves the problems of digital modeling of urban environments. Recently new approaches for this purpose have been developed with the idea of using the information from different sensors, depth laser sensors, CCD cameras and inertial systems primarily [9] [4].

Information fusion from multiple sources is an important issue when processing data as it must have an accurate calibration of all the instruments of the acquisition platform, a good real-time data fusion technique is proposed in [9], they fuse colour camera, range scanning anda navigation data to produce a three dimensional colour point clouds. For this article we will focus on characterizing

J.A. Carrasco-Ochoa et al. (Eds.): MCPR 2013, LNCS 7914, pp. 104–113, 2013.

the rigid transformation between the camera and laser sensor. Using the intrinsic parameters of the camera and LIDAR, and the transformation we project a 3D point from LIDAR to the image plane of the camera. The problem of rigid transformation is solved by matching features of the 3D calibration pattern acquired from LIDAR and camera. These features are usually the corners of planar checkboard pattern, in the image as an intersection of black and white square, in the LIDAR correspont the intersection to square plane and hole square.

In this work, two types of sensors are used, *Velodyne HDL-64E* laser scanner and *Point Grey Ladybug2* spherical digital camera. The laser scanner operates by pulsing a laser diode for a short duration (typically 4 nanoseconds) and precisely measuring the amount of time it takes for the pulse to travel to an object, reflect off and return to a highly sensitive photodetector and camera system is a high resolution omnidirectional sensor, it has six $0.8 - Megapixel$ cameras, with five CCDs positioned in a horizontal ring and one positioned vertically, that enable the system to collect video from more than 75% of the full sphere. Complementary benefits of both sensors, it's possible to acquire more reliable scenes have characteristics such as depth, color and orientation. Both sensors are georeferenced. Data fusion approaches between LIDAR and images can be divided in two categories [11]:

- Centralized: the data fusion process occurs at pixel or feature level, i.e., the LIDAR and camera characteristics are combined into a single vector for subsequent classification. One drawback to this approach is that only areas that are seen by both sensors can be processed.
- Decentralized: data processing from each sensor is made separately and then fusion occurs. These methods usually require training to determine the fusion model.

The estimation of LIDAR and camera intrinsic parameters is a nonlinear problem that can be solved in different ways. A novel algorithm is proposed [5] for joint estimation of both the intrinsic parameters of the laser sensor and the LIDAR-camera transformation. Specifically, they use measurements of a calibration plane at various configurations to establish geometric constraints between the LIDAR's intrinsic parameters and the LIDAR-camera 6 d.o.f. relative transformation. They process these measurement constraints to estimate the calibration parameters as follows: First, analytically compute an initial estimate for the intrinsic and extrinsic calibration parameters in two steps. Subsequently, they employ a batch iterative (nonlinear) least-squares method to refine the accuracy of the estimated parameters. Other method using a planar board checkerboard patterns is proposed in [6], the autors defines the rotation matrix between the sensor laser and camera as achieved moving they platform and observing the resulting motion of the sensors, this step attempts to solve the well-known homogeneous transform, translation is calculated using a commonly least-squares estimation algorithm according to the corners of the pattern, detected by both sensors. Besides the problem of calibration in [8] provide a solution for the occlusion problem that arises in conjuntion with different view points of the fusioned sensors, they approach first perfoms a synchronization of both sensors to allow for moving objects which incorporates an inertial correction of the LIDAR data and an automatic trigger mechanism depending on the view direction of the camera, the occlusion detection algorithm uses 2D convex hulls derived from a 3D object segmentation to allow for a visibility check of projected 3D LIDAR data.

Not only square flat patterns are used for calibration between these two sensors, in [7] they use a circle-based calibration object because its geometry allows to obtain an accurate estimation pose in the camera frame and the camera intrisic parameters. The autors use a linear minimization of the Euclidian distance error between the 3D circle center point sets, then they first generate the 3D circles of the n poses estimated by the camera, it consists in computing m points of each estimated circle pose by using the 3D circle center and an orthonormal base lying in circle's plane. Other approach using a conic based geometry object to calibrate 2D/3D laser sensors is presented in [1].

2 Platform Projection Model

The goal of the method is to find homogeneous transformation between the pinhole camera and the LIDAR in order to fuse the measurements from both sensors in urban digitalize enviroments applications.

Data is collected by a mobil platform using the data collection vehicle shown in Figure 1. This mobile mapping system is composed of a LIDAR sensor, video camera, GPS, Inertial Measurement Unit (IMU).

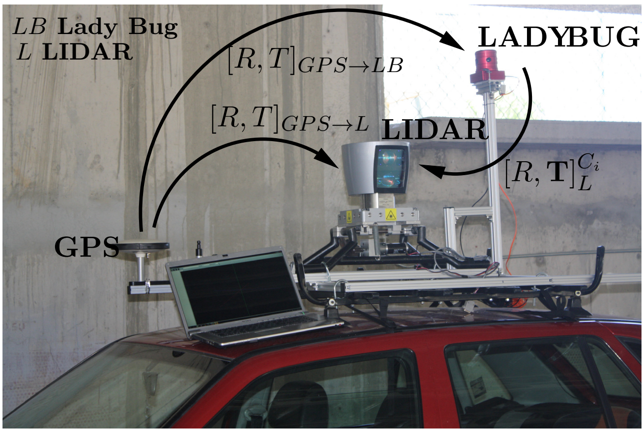


Fig. 1. Sensor platform composed of LIDAR Velodyne HDL-64E (L) Ladybug2 (LB) and GPS. Ladybug2® spherical digital video camera system has six cameras (C_i). L, C_i represent the LIDAR and six camera frames of the Ladybug.

This research work focuses the refinement of the registration of LIDAR and panoramic images.

2.1 Panoramic Camera and LIDAR Model

The six cameras frames of the *Ladybug2* are represented by C_i (where $i = 1, \ldots, 6$). The 3D points acquired by the LIDAR ($\mathbf{X}^L$) are transform from LIDAR frame to camera frames by $[R_L^{C_i}, \mathbf{T}_L^{C_i}]$, called the extrinsic parameters (see Figure 2.)

A 2D point in the camera C_i is denoted by $\mathbf{u}^{C_i} = \begin{bmatrix} u^{C_i} \; v^{C_i} \end{bmatrix}^T$. A 3D point in LIDAR frame is denoted by $\mathbf{X}^L = \begin{bmatrix} X^L \; Y^L \; Z^L \end{bmatrix}^T$. We use $\hat{\mathbf{x}}$ to denote the augmented vector by adding 1 as the last element: $\hat{\mathbf{u}}^{C_i} = \begin{bmatrix} u^{C_i} \; v^{C_i} \; 1 \end{bmatrix}^T$ and $\hat{\mathbf{X}}^L = \begin{bmatrix} X^L \; Y^L \; Z^L \; 1 \end{bmatrix}^T$. A camera is modeled by the usual pinhole: The image $\mathbf{u}^{C_i}$ of a 3D point $\mathbf{X}^L$ is formed by an optical ray from $\mathbf{X}^L$ passing through the optical center C_i and intersecting the image plane. The relationship between the 3D point $\mathbf{X}^L$ and its image projection $\mathbf{u}^{C_i}$ is given by

$$s\hat{\mathbf{u}}^{C_i} = A^{C_i}[R_L^{C_i}, \mathbf{T}_L^{C_i}]\hat{\mathbf{X}}^L = P_L^{C_i}\hat{\mathbf{X}}^L \tag{1}$$

with $A^{C_i} = \begin{bmatrix} -k_u\, f & 0 & u_0 & 0 \\ 0 & k_v\, f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}^{C_i}$

and $P = A^{C_i}[R_L^{C_i}, \mathbf{T}_L^{C_i}]$
where s is an arbitrary scale factor, $[R_L^{C_i}, \mathbf{T}_L^{C_i}]$, called the extrinsic parameters, is the rotation and translation which relates from LIDAR system L to camera system C_i, and A^{C_i} is called intrinsic matrix for the camera C_i, with (u_0, v_0) the coordinates of the principal point, fk_u and fk_v the scale factors in image u and v axes. The 3×4 matrix $P_L^{C_i}$ is called the camera projection matrix, which mixes both intrinsic and extrinsic parameters.

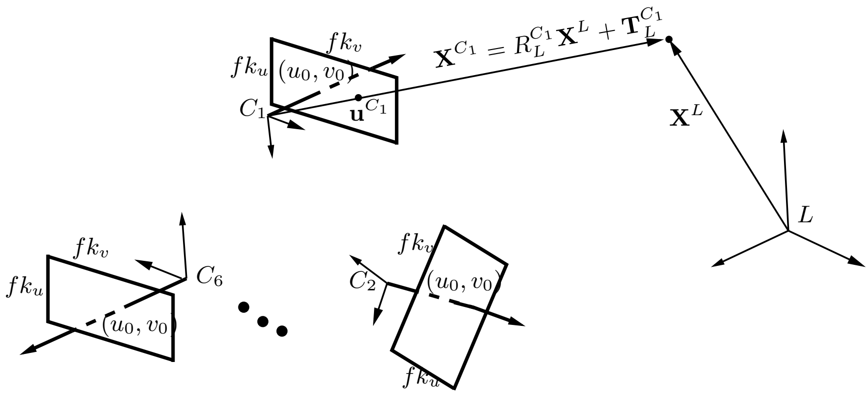


Fig. 2. LIDAR frame L and the six pinhole camera model

3 Panoramic Camera and LIDAR Calibration

The six cameras of the Ladybug2 are calibrated with the Zhang's method [10], in this the intrinsic parameters (A^{C_i}) of the cameras are computed. The LIDAR is calibrated using the method proposed in [2]. We use a pattern which facilitates the extraction of the same point in camera and LIDAR data. This pattern is shown in Figure 7 and it is composed of white rhombus and rhombus holes which take the black color by the background. The Figure 3 shows the relationship between the pattern W_i camera C_i and LIDAR L. The transformation between LIDAR and cameras shown in te Equation 1 is computed by

$$[R_L^{C_i}, \mathbf{T}_L^{C_i}] = [R, \mathbf{T}]_W^{C_i} * ([R, \mathbf{T}]_W^L)^{-1} \tag{2}$$

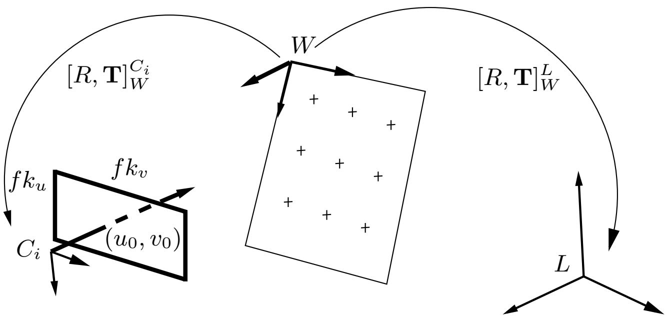


Fig. 3. The extrinsic parameters $[R, \mathbf{T}]_W^{C_i}$ are computed by Zhang's method and $[R, \mathbf{T}]_W^L$ are computed by Algorithm shows in Figure 4

In the algorithm shows in Figure 4, the function $transform\_to\_XYZ$ transform the LIDAR data to X, Y, Z points using the method proposed in [2]. RANdom SAmple Consensus ($RANSAC$) algorithm is widely used for plane detection in point cloud data. The principle of $RANSAC$ algorithm consists to search the best plane among a 3D point cloud. For this purpose, it selects randomly three points and it calculates the parameters of the corresponding plane. Then it detects all points ($nPatternPointCloud$) of the original cloud belonging to the calculated plane (Π), according to a given threshold. The $project\_to\_normal\_pattern$ algorithm projected the points $nPatternPointCloud$ to plane Π. In function $MaximizingFunction$, we build an artificial calibration pattern ($Pattern(radius, \theta)$) using the dimension of the pattern calibration. The artificial pattern can be moved on the plane Π rotating an angle θ and moving a distance $radius$ in the direction of rotation. The $nProjectedPattern$ points are comparing with the artificial pattern plane, the maximum comparison allows us to calculate the distance ($radius$) and angle (θ) in the plane where the reference pattern placed. Using the plane equation Π and ($radius, \theta$) we computing the $[R, \mathbf{T}]_W^L$ parameters.

For each camera C_i do:

Algorithm for computing of extrinsic LIDAR parameters.

for each nRawPointCloud **do**
$nPointCloud \leftarrow transform\_to\_XYZ(nRawPointCloud)$
$(nPatternPointCloud, \Pi) \leftarrow RANSAC(nPointCloud)$ ▷ extract pattern points
$nProjectedPattern \leftarrow project\_to\_normal\_pattern(nPatternPointCloud)$
$[R, \mathbf{T}]_W^L \leftarrow MaximizingFunction(nProjectedPattern, \Pi)$
end for

Algorithm for computing extrinsic and intrinsic *Ladybug2* parameters.

for each nImage **do**
$[R, \mathbf{T}]_W^{C_i}, A^{C_i} \leftarrow$ `use Bouguet's camera calibration Toolbox`
end for

LIDAR and Ladybug calibration

for each nImage **do**
$[R_L^{C_i}, \mathbf{T}_L^{C_i}] = [R, \mathbf{T}]_W^{C_i} * ([R, \mathbf{T}]_W^L)^{-1}$
end for

Fig. 4. Algorithms used in this work

4 Results

The results are presented in three stages, the computation of extrinsic parameters of the LIDAR, the computation of extrinsic and intrinsic parameters of the cameras C_i of the Ladybug and extrinsic parameters between each camera and LIDAR. For practical purposes, we show the results by only one camera, however the Ladybug is formed with six cameras.

4.1 Algorithm for Computing of Extrinsic LIDAR Parameters

The Figure 5 shows the exracted points (blue points) using the $RANSAC$ algorithm and they proyected onto plane Π using the $project\_to\_normal\_pattern$ algorithm. The red circle in the Figure 5 represent the position of the artificial pattern plane. This position is compute by the algorithm $MaximizingFunction$. The $MaximizingFunction$ computes the matches between the real pattern data and the synthetic pattern, this is shown in Figure 6. The rotation and translation are performed on the plane Π, and the traslation is carried out in θ orientation.

The rigid transformation of a 3D point in the LIDAR frame, L, into the world frame (pattern frame) is defined by the rotation matrix and translation vector ($[R, \mathbf{T}]_W^L$):

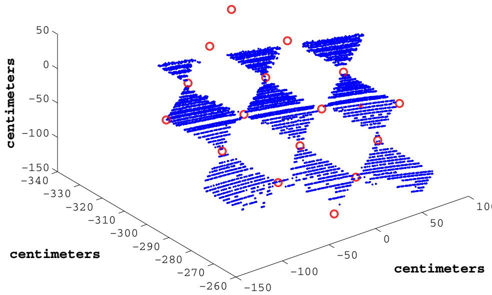


Fig. 5. Extraction of the calibration pattern using the *RANSAC* algorithm and projection onto the plane Π

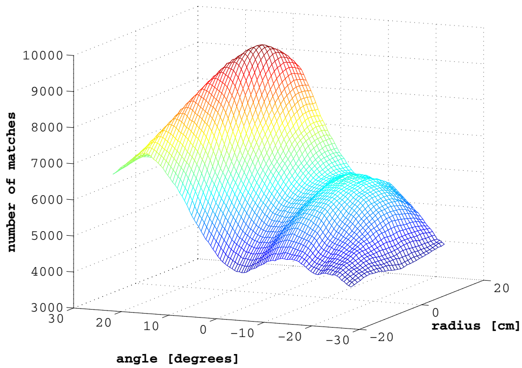


Fig. 6. The maximization function computes the number of points of the real pattern matches with the synthetic pattern

$$R_W^L = \begin{bmatrix} 0.1085 & -0.3006 & 0.9475 \\ -0.7800 & -0.6167 & -0.1063 \\ -0.6164 & 0.7275 & 0.3014 \end{bmatrix}; \mathbf{T}_W^L = \begin{bmatrix} -282.90 \\ 80.71 \\ -49.82 \end{bmatrix}$$

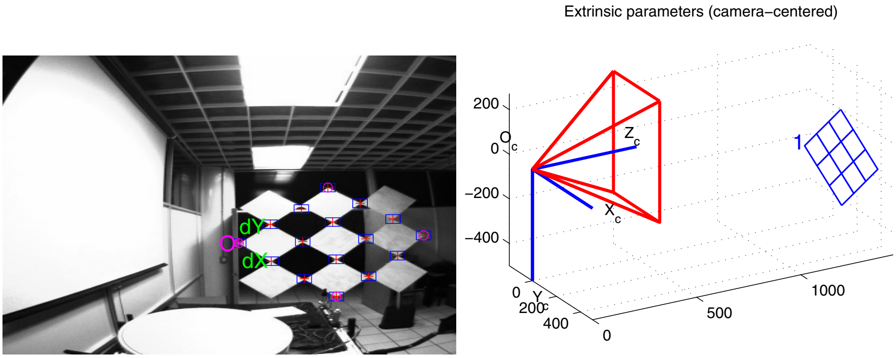
4.2 Algorithm for Computing Extrinsic and Intrinsic *Ladybug* Parameters

We use camera calibration toolbox for Matlab [3]. The intrinsic camera parameters are:

intrinsic parameters	value
Focal length	(628.4651, 622.5191) px
Principal point location	(348.0818, 507.8548) px
Distortion coefficients	(-0.3759,0.1139,0.0027, 0.0049,0) px^2
Skew coefficient	0

The Figure 7(a) shows the subpixelic extraction points. The pattern plane position is the same for the pattern plane acquired by the LIDAR in the Figure 5. The world frame is shown in the Figure 7(b) with respect to the camera frame. The $[R, \mathbf{T}]_W^{C_i}$ are:

$$R_W^{C_i} = \begin{bmatrix} 0.6588 & 0.6417 & -0.3926 \\ 0.6257 & -0.7572 & -0.1877 \\ -0.4178 & -0.1220 & -0.9003 \end{bmatrix} ; \mathbf{T}_W^{C_i} = \begin{bmatrix} 35.0 \\ 76.35 \\ 365.36 \end{bmatrix}$$



(a) (b)

Fig. 7. (a) Subpixel corners extration. (b) World frame respect to the camera frame.

4.3 LIDAR and Ladybug Calibration

The rigid transformation between camera and LIDAR frame is computed using the Equation 2. The result is shown in Figure 8. In this figure, the pattern acquired by the LIDAR is transformed onto the image frame using the extrinsic parameters $[R_L^{C_i}, \mathbf{T}_L^{C_i}]$, this transformation allows us referenced in the camera the points acquired by the LIDAR. The projection is completed using the intrinsic camera parameters. The Figure 9 demostrate the calibration system. The Figure 9 (a) shows the LIDAR data that correspond a person, the 3D data are projected into the image. The projected points are show in Figure 9 (b) as red points.

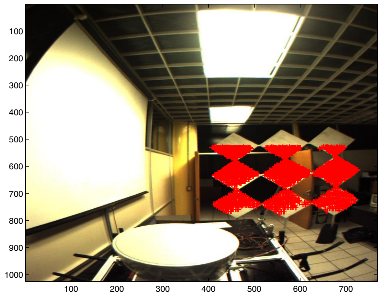
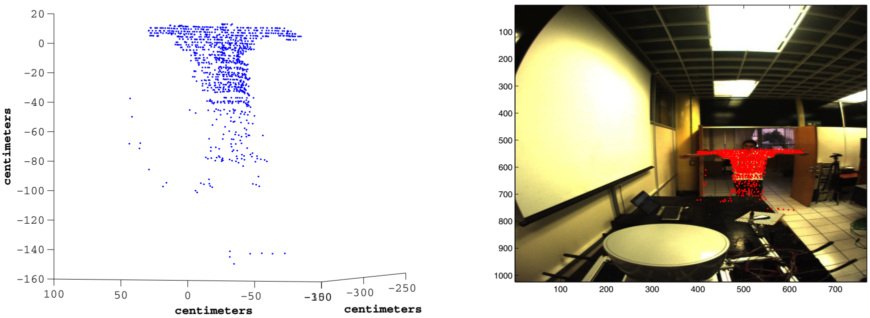


Fig. 8. The red points are acquired by the LIDAR and projected to image



(a) (b)

Fig. 9. (a) A person acquired by the LIDAR. (b) The red points correspond the 3D points projected into image.

5 Conclusions and Future Work

A new extrinsic calibration method for a multiple sensor has been proposed. By using a plane calibration target, extrinsic calibration between sensors and intrinsic camera calibration can be effectuated simultaneously. The results obtained in real data tests illustrate an appropriate and accurate projection of the lidar data. The future work are oriented to estimate the confidence intervals in the calibration method and the error propagation in data sensor fusion for texture 3D data.

References

1. Almeida, M., Dias, P., Oliveira, M., Santos, V.: 3d-2d laser range finder calibration using a conic based geometry shape. In: Image Analysis and Recognition, pp. 312–319 (2012)

2. Atanacio-Jiménez, G., González-Barbosa, J.J., Hurtado-Ramos, J.B., Ornelas-Rodríguez, F.J., Jiménez-Hernández, T., García-Ramirez, H., González-Barbosa, R.: Lidar velodyne hdl-64e calibration using pattern planes. International Journal of Advanced Robotic Systems 8(5), 70–82 (2011)
3. Bouguet, J.Y.: Camera calibration toolbox for Matlab (2008)
4. Mei, C., Rives, P.: Calibration between a central catadioptric camera and a laser range finder for robotic applications. In: Proceedings 2006 IEEE International Conference on Robotics and Automation, ICRA 2006, pp. 532–537. IEEE (2006)
5. Mirzaei, F.M., Kottas, D.G., Roumeliotis, S.I.: 3d lidar-camera intrinsic and extrinsic calibration: Observability analysis and analytical least squares-based initialization. International Journal of Robotics Research 31(4), 452–467 (2012)
6. Nunez, P., Drews Jr, P., Rocha, R., Dias, J.: Data fusion calibration for a 3d laser range finder and a camera using inertial data. In: European Conference on Mobile Robots, vol. 9, p. 9 (2009)
7. Rodriguez, F., Fremont, V., Bonnifait, P., et al.: Extrinsic calibration between a multi-layer lidar and a camera. In: IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, MFI 2008, pp. 214–219. IEEE (2008)
8. Schneider, S., Himmelsbach, M., Luettel, T., Wuensche, H.J.: Fusing vision and lidar-synchronization, correction and occlusion reasoning. In: 2010 IEEE Intelligent Vehicles Symposium (IV), pp. 388–393 (2010)
9. Underwood, J., Scheding, S., Ramos, F.: Real-time map building with uncertainty using colour camera and scanning laser. In: Proceedings of the 2007 Australasian Conference on Robotics and Automation (2007)
10. Zhang, Z.: A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1330–1334 (2000)
11. Zhao, G., Xiao, X., Yuan, J.: Fusion of velodyne and camera data for scene parsing. In: 2012 15th International Conference on Information Fusion (FUSION), pp. 1172–1179 (2012)