Extrinsic Calibration of a 3D Laser Scanner and an Omnidirectional Camera ⋆

Gaurav Pandey * **James McBride** ** **Silvio Savarese** *
Ryan Eustice *

* *University of Michigan, Ann Arbor, MI 48109 USA (e-mail: pgaurav,eustice,silvio@umich.edu)*
** *Ford Motor Company Research and Innovation Center, Dearborn, MI 48124 USA (e-mail: jmcbride@ford.com)*

Abstract: We propose an approach for external calibration of a 3D laser scanner with an omnidirectional camera system. The utility of an accurate calibration is that it allows for precise co-registration between the camera imagery and the 3D point cloud. This association can be used to enhance various state of the art algorithms in computer vision and robotics. The extrinsic calibration technique used here is similar to the calibration of a 2D laser range finder and a single camera as proposed by Zhang (2004), but has been extended to the case where we have a 3D laser scanner and an omnidirectional camera system. The procedure requires a planar checkerboard pattern to be observed simultaneously from the laser scanner and the camera system from a minimum of 3 views. The normal of the planar surface and 3D points lying on the surface constrain the relative position and orientation of the laser scanner and the omnidirectional camera system. These constraints can be used to form a non-linear optimization problem that is solved for the extrinsic calibration parameters and the covariance associated with the estimated parameters. Results are presented for a real world data set collected by a vehicle mounted with a 3D laser scanner and an omnidirectional camera system.

Keywords: Sensor Calibration, 3D Laser Scanner, Omnidirectional Camera.

1. INTRODUCTION

One of the basic tasks of mobile robotics is to automatically create a 3D map of the environment. However, to create realistic 3D maps, we need to acquire visual information (e.g. color, texture) from the environment and this information has to be precisely mapped onto the range information. To accomplish this task, the camera and 3D laser range finder must be extrinsically calibrated, i.e., the rigid body transformation between the two reference systems must be estimated. This rigid body transformation allows reprojection of the 3D points from the laser coordinate frame to the 2D coordinate frame of the image. Most previous works on extrinsic laser-camera calibration concern calibration of perspective cameras to 2D laser scanners (Zhang (2004)). Mei and Rives (2006) have described the calibration of a 2D laser range finder and an omnidirectional camera. They showed the results for both visible (laser is observed in camera image also) and invisible lasers. Unnikrishnan and Hebert (2005) extended Zhang's (Zhang (2004)) method to calibrate a 3D laser scanner with a perspective camera. Recently Aliakbarpour et al. (2009) have proposed a novel approach for calibration of a 3D laser scanner and a stereo camera, which uses an Inertial Measurement Unit (IMU) to decrease the number of points needed for a robust calibration.
In contrast to previous works, here we consider the extrinsic calibration of an omnidirectional camera with a 3D laser range finder. The problem of extrinsic calibration of a 3D scanner and an omnidirectional camera was first addressed by Scaramuzza et al. (2007). There, they proposed a technique that requires manual selection of point correspondences from a scene viewed from the two sensors. In this work, we describe a method of extrinsic calibration of an omnidirectional camera and a high resolution 3D laser scanner (with invisible lasers) that does not require any explicit point correspondence.
The outline of the paper is as follows: In Section 2.1 we describe a procedure for the automatic refinement of the intrinsic calibration of the Velodyne laser scanner range correction. Section 2.2 describes the omnidirectional camera system used. Section 2.3 describes the proposed extrinsic laser-camera calibration method and in Section 3 we present some calibration results. In Section 4 we discuss the implications of the laser-camera mapping presented in this paper.

2. METHODOLOGY

Extrinsic calibration requires co-observable features in both camera and laser data, moreover these features should be easy to extract from both sensor modalities. In our calibration procedure we employ a checkerboard pattern mounted on a planar surface, which we will refer to as the target plane from now onwards. Our selection of the target is based on (i) the checkerboard pattern is easy to extract from the image data; and (ii) its planar structure is easy to extract from the 3D point cloud. The

⋆ This work is supported through a grant from Ford Motor Company via the Ford-UofM Alliance (Award #N008265).

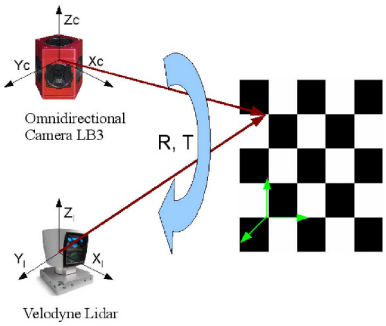


Fig. 1. Depiction of the experimental setup used for extrinsic calibration of the 3D laser scanner with omnidirectional camera system.

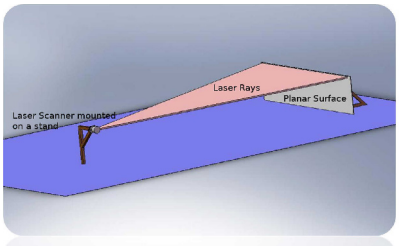


Fig. 2. Depiction of the manual calibration technique used by the manufacturers to estimate the range offset corresponding to each laser.*[Figure courtesy: Velodyne manufacturer]*

approximate setup for extrinsic calibration along with the perception sensors (Velodyne HDL-64E 3D laser scanner and Pointgrey Ladybug3 omnidirectional camera) used for the experiments is depicted in Fig. 1.

2.1 Velodyne Laser Scanner

The Velodyne HDL-64E is a high definition lidar (HDL) sensor designed to meet the demands of autonomous navigation, surveying, mapping and other applications. This commercially available sensor became popular in the 2007 DARPA Urban Challenge and since then it has been used in various research works. The HDL-64E 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. This results in a time of flight (TOF) range measurement D_l for each pulsed laser.
The HDL-64E has two blocks of lasers each consisting of 32 laser diodes mounted on the front assembly with photo-detectors in the middle. Each laser diode is precisely aligned at predetermined vertical angles, resulting in an effective 26.8 degree vertical field of view. The entire unit can spin about its vertical axis at speeds up to 900 rpm (15 Hz) to provide a full 360 degree azimuthal field of view. More technical details about the sensor can be found in McBride et al. (2008). The calculated range measurement D_l contains some bias due to the errors in TOF calculation and thus has a bias correction δD from the actual measurement. The manufacturers of the HDL-64E provide a per laser calibration value of δD to compensate the error in range measurement. They obtain an approximate value of this δD correction by

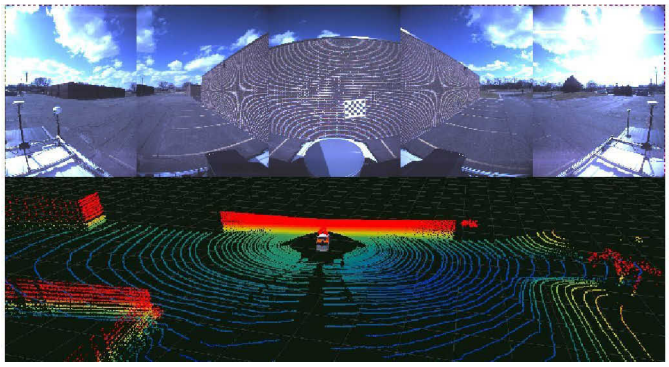


Fig. 3. Depiction of the proposed automatic calibration procedure. The sensor is placed in front of a wall and laser measurements are recorded. The omnidirectional image is for visualization only and not required for the calibration procedure.

manually calibrating each laser range measurement using a calibration procedure as depicted in Fig. 2. The laser scanner is mounted on a support in front of a wall and the offset in the range measurement is calculated by considering the manually measured distance between the wall and the laser scanner as ground truth, i.e.

$$\delta D = D_m - D_l, \tag{1}$$

where D_m is the range calculated manually by measuring the distance between the wall and the sensor and D_l is the TOF range measurement.
The extrinsic calibration of the laser-camera sensors is greatly affected by the intrinsics of the sensors themselves. Therefore, it is important that the laser scanner is well calibrated, so before considering the problem of extrinsic calibration we propose a robust way to automatically calculate this optimum offset δD in the range measurement. In contrast to the calibration method used by the manufacturers we propose an in-situ method that only requires the user to bring the platform mounted with the laser scanner in front of a wall or a planar surface and record the range measurements (Fig. 3). The laser measurements are recorded for different positions of the sensor platform in front of the wall or planar surface. Now if we use the δD correction as calculated in (1) and consider the points lying on the wall, they should all be coplanar. But since the δD corrections are not exact the reprojection error of these points on to the estimated plane is significant. We can thus minimize this reprojection error over different values of δD, jointly for all the lasers, to get an optimum value of the individual range corrections. We use RANSAC (Fischler and Bolles (1981)) to estimate the equation of the best fit plane for all the points lying on the wall. We first generate a bounding box that contains the target plane and establish potential laser points lying in the box. Then, these potential laser points $\{\tilde{Q}_l^i; i = 1, 2, \cdots, N\}$ are passed to the RANSAC plane fitting algorithm, which returns the set of inliers (3D points in laser reference frame) corresponding to the best fit plane to these potential laser points. The RANSAC plane fitting algorithm can be described in the following steps:

(1) Randomly choose 3 points from $\{\tilde{Q}_l^i; i = 1, 2, \cdots, N\}$.
(2) Find the equation of the plane passing through these points.

(3) Find the inliers corresponding to the plane calculated in step 2.
(4) Repeat until we find the best plane, i.e., the plane containing most of the points.

Let the plane calculated by RANSAC be parametrized by the normal to the plane from the origin of the laser coordinate frame, given by $\mathbf{N} = [n_x, n_y, n_z]^\top$, such that $\|\mathbf{N}\|$ is the perpendicular distance of the plane from the origin. So if $\tilde{P} = [X, Y, Z]^\top$ is any point lying in this plane then the projection of the vector $\mathbf{P}$ (from origin to point P) on the normal $\mathbf{N}$ is equal to the length of the normal itself i.e.,

$$\mathbf{P} \cdot \mathbf{N} = ||\mathbf{N}||^2. \tag{2}$$

If D be the range of this point, measured by laser i, and θ and ω be the corresponding elevation and azimuth angle, respectively, then:

$$X = D\cos\theta\sin\omega, \tag{3}$$
$$Y = D\cos\theta\cos\omega, \tag{4}$$
$$Z = D\sin\theta. \tag{5}$$

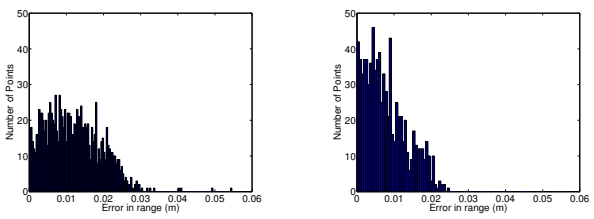
So the actual range of this point in terms of the plane normal can be written as:

$$D = \frac{\|\mathbf{N}\|^2}{n_x\cos\theta\sin\omega + n_y\cos\theta\cos\omega + n_z\sin\theta}. \tag{6}$$

So now we have the range measurement as obtained from the RANSAC plane fitting algorithm and we have one range measurement from the manufacturers, $D_m = D_l + \delta D$, where δD is the distance correction calculated by manual calibration. Here we find the offset $\delta D'$, which when added to the laser range measurement, projects the laser point to the wall or the calculated plane. This is obtained by iteratively minimizing the following non linear least squares error for all 64 lasers and all the points lying on the plane:

$$\delta D_i{}' = \underset{\delta D_i{}'}{\operatorname{argmin}} \sum_{i=1}^{64}\sum_{j=1}^{n} ||D_{ij} - (D_{l_{ij}} + \delta D_i{}')||^2, \tag{7}$$

where D_{ij} is the range of the jth point corresponding to ith laser as calculated from (6) and $D_{l_{ij}}$ is the corresponding TOF range measurement from the laser. The result of the above mentioned optimization process is shown in Fig. 4.



(a) Before Correction (b) After Correction

Fig. 4. Histogram of error in range of points falling on the plane. The error is the perpendicular distance of the inlier points from the estimated plane. The plane equation and the inliers are estimated using RANSAC.

2.2 Ladybug3 Omnidirectional Camera

The Pointgrey Ladybug3 (LB3) is a high resolution omnidirectional camera system. It has six 2-Megapixel cameras, with five CCDs positioned in a horizontal ring and one positioned vertically, that enable the system to collect video from more than 80% of the full sphere. More technical details about the camera can be obtained from the manufacturer's website (Pointgrey (2009)). The camera is pre-calibrated from the manufacturer so that the intrinsic parameters of individual camera are well known. Moreover, the rigid body transformation of all the cameras with respect to a common coordinate frame called the camera head is also known. Therefore, we need to estimate the pose (orientation and position) of the camera head (with respect to some local reference frame) so that we can represent any 3D point in the camera head's frame and thereafter to the coordinate frame of any camera. We used Zhang's (Zhang (1998)) method to calculate pose of the camera head with respect to the local reference frame attached to the target plane as discussed in section 2.3.

2.3 Extrinsic Calibration of 3D Laser Scanner and the Omnidirectional Camera System

The extrinsic calibration technique is similar to the one proposed by Zhang (2004), which requires the system to observe a planar pattern in several poses, and the constraints are based upon data captured simultaneously from the camera and the laser scanner. The normal to the target plane and the laser points on the target plane are related, and constrain the relative position and orientation of the camera and laser scanner. We know the equation of the target plane in the coordinate system attached to the plane itself, which for convenience is given by:

$$Z = 0. \tag{8}$$

Let $\tilde{P}_w$ be the coordinate of any point in the world reference frame (here it is the coordinate frame attached to the target plane) and ${}^{c_i}_{w}R$ be the orthonormal rotation matrix that rotates frame w (world frame) into frame c_i (ith camera frame) and ${}^{\mathbf{c_i}}\mathbf{t}_{\mathbf{c_i w}}$ be the Euclidean 3-vector from c_i to w as expressed in frame c_i. Then the transformation equation that transforms a point from the world reference frame to the reference frame of the ith camera can be written as:

$$\tilde{P}_{c_i} = {}^{c_i}_{w}R\tilde{P}_w + {}^{\mathbf{c_i}}\mathbf{t}_{\mathbf{c_i w}}, \tag{9}$$

where $\tilde{P}_{c_i}$ is the coordinate of that same point in ith camera's reference frame. Since we know the transformation matrices ${}^{h}_{c_i}R$ and ${}^{\mathbf{h}}\mathbf{t}_{\mathbf{hc_i}}$ that transform a point from the ith camera frame to the camera head frame, we can write the coordinate of this point in the camera head frame as:

$$\tilde{P}_h = {}^{h}_{c_i}R\tilde{P}_{c_i} + {}^{\mathbf{h}}\mathbf{t}_{\mathbf{hc_i}}. \tag{10}$$

Thus, we can transform any point $\tilde{P}_w$, lying in the target plane, into the camera head reference frame if we know the transformation ${}^{c_i}_{w}R$ and ${}^{\mathbf{c_i}}\mathbf{t}_{\mathbf{c_i w}}$. We used Zhang's (Zhang (1998)) method for finding this transformation relative to the planar target.

For a usual pin hole camera model, the relationship between a homogeneous 3D point $\tilde{P}_w = [X\ Y\ Z\ 1]^\top$ and its image projection $\tilde{p} = [u\ v\ 1]^\top$ is given by:

$$\tilde{p} = K_i[{}^{c_i}_{w}R\ {}^{\mathbf{c_i}}\mathbf{t}_{\mathbf{c_i w}}]\tilde{P}_w, \tag{11}$$

where $({}^{c_i}_{w}R, {}^{\mathbf{c_i}}\mathbf{t}_{\mathbf{c_i w}})$, called the extrinsic parameters, are the rotation and translation that relates the world coordinate system to the camera coordinate system, and K_i is the camera intrinsic matrix.
Assuming that the image points are corrupted by independent and identically distributed noise, the maximum likelihood estimate of the required transformation $({}^{c_i}_{w}R, {}^{\mathbf{c_i}}\mathbf{t}_{\mathbf{c_i w}})$ can be obtained by minimizing the following reprojection error (Zhang (1998)) for n images of the target plane and m points per image:

$$\underset{{}^{c_i}_{w}R, {}^{\mathbf{c_i}}\mathbf{t}_{\mathbf{c_i w}}}{\operatorname{argmin}} \sum_{k=1}^{n}\sum_{j=1}^{m} \|\tilde{p_{kj}} - K_i[{}^{c_i}_{w}R \; {}^{\mathbf{c_i}}\mathbf{t}_{\mathbf{c_i w}}]\tilde{P}_j\|. \quad (12)$$

Here, ${}^{c_i}_{w}R$ is an orthonormal rotation matrix parametrized by the 3 Euler angles. Now, if ${}^{c_i}_{w}R = [\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}]$ and ${}^{\mathbf{c_i}}\mathbf{t}_{\mathbf{c_i w}}$ is the Euclidean 3-vector from c_i to w as expressed in frame c_i then we can write the equation of the target plane in the ith camera frame as:

$$\mathbf{r_3} \cdot (\mathbf{p} - {}^{\mathbf{c_i}}\mathbf{t}_{\mathbf{c_i w}}) = 0, \quad (13)$$

where $\mathbf{p}$ is the vector from origin to any point lying on the plane.

Therefore, the normal of the target plane in the ith camera frame is given by:

$$\mathbf{N_{c_i}} = (\mathbf{r_3} \cdot {}^{\mathbf{c_i}}\mathbf{t}_{\mathbf{c_i w}})\mathbf{r_3}. \quad (14)$$

Here, $\|\mathbf{N_{c_i}}\| = \mathbf{r_3} \cdot {}^{\mathbf{c_i}}\mathbf{t}_{\mathbf{c_i w}}$ is the distance of the target plane from the ith camera's center. Since we know the pose of the ith camera with respect to the camera head we can calculate the normal of the plane N_h in the camera head frame.

$$\mathbf{N_h} = {}^{h}_{c_i}R\mathbf{N_{c_i}} \left(\|\mathbf{N_{c_i}}\| + \frac{\mathbf{N_{c_i}} \cdot {}^{\mathbf{h}}\mathbf{t}_{\mathbf{hc_i}}}{\|\mathbf{N_{c_i}}\|} \right). \quad (15)$$

Once we know the normal vector to the target plane in camera head's reference frame, we need to find the 3D points in the laser reference frame that lie on the target plane. We use the RANSAC plane fitting algorithm described in section 2.1 to compute these 3D points. We also know the normal vector to the target plane from (15). These two measures provide a constraint on the required 3D rigid body transformation between the laser and the camera system. Let $\{\tilde{P}_l^i; i = 1, 2, \cdots, n\}$ be the set of 3D points lying on the plane given by RANSAC; the coordinates of these points are known in the laser reference system. The coordinates of these points in the camera head's frame are given by:

$$\tilde{P}_h^i = {}^{h}_{l}R\tilde{P}_l^i + {}^{\mathbf{h}}\mathbf{t}_{\mathbf{hl}}, \quad (16)$$

where ${}^{h}_{l}R$ and ${}^{\mathbf{h}}\mathbf{t}_{\mathbf{hl}}$ are the required rotation and translation matrices that project any point in the laser reference system to the camera head's frame and thereby to the respective camera. Now, if we shoot a ray from the camera head to any point $\tilde{P}_h^i$ lying on the plane, the projection of this ray on to the normal of the plane is equal to the distance of the plane from the origin. Therefore for m different views of the target plane and n 3D laser points per view, the laser-camera extrinsic parameters can be obtained by minimizing the following reprojection error:

$$F = \sum_{i=1}^{m}\sum_{j=1}^{n} \left(\frac{\mathbf{N_h^i}}{\|\mathbf{N_h^i}\|} \cdot ({}^{h}_{l}R\mathbf{P}_l^j + {}^{\mathbf{h}}\mathbf{t}_{\mathbf{hl}}) - \|\mathbf{N_h^i}\| \right)^2, \quad (17)$$

where $\mathbf{N_h^i}$ is the normal to the i^{th} pose of the target plane in the camera head's frame. We can solve the non linear optimization problem given in (17) for ${}^{h}_{l}R$ and ${}^{\mathbf{h}}\mathbf{t}_{\mathbf{hl}}$ using Levenberg Marquadrt algorithm (Levenberg (1944), Marquadrt (1963)).

2.4 Minimum number of views required

Fig. 5. Geometrical interpretation of minimum number of views required for calibration. (a) The translation of the sensors along the target plane and rotation about the axis parallel to normal of the plane is not constrained. (b) The translation of the sensors along the line of intersection of the two planes is not constrained.

A minimum of three non-coplanar views of the target plane are required to fully constrain the optimization problem (17) for the estimation of the calibration parameters. If only one plane is considered, as shown in Fig. 5(a), then the cost function (17) does not change when the sensors are either translated along the plane parallel to the target plane or rotated about the axis parallel to the normal of the target plane. Thus the solution obtained from a single view does not converge to the actual value in the following three parameters: 2D Translation along the target plane and a rotation about the normal of the target plane. Similarly for two views (Fig. 5(b)) the translation of the sensor along the line of intersection of the two planes does not change the cost function, thereby giving large uncertainity in that direction. Three views are required to completely constrain the 6 degee of freedom (DOF) pose of one sensor with respect to the other.

2.5 Covariance of the estimated parameters

The parameters estimated by minimizing the cost function given in (17) have some error due to the uncertainity in the sensor measurements. The laser we have used in our experiments has unertainity in the range measurements of the order of $0.02m$. This uncertainity due to the random perturbations of the range measurements is propagated to the estimated parameters. It is very important to know this uncertainity in order to use the parameters calculated here in any vision or SLAM algorithm. Haralick (1998) has described a method to propagate the covariance of the measurements through any kind of scalar non-linear optimization function. The only assumptions are that the scalar function be non-negative, has finite first and second order partial derivatives, that its value be zero for ideal data, and the random perturbations in the input be small enough so that the output can be approximated by the first order Taylor series expansion. The optimization function (17) we use here satisfies these assumptions, so we can calculate the covariance of the estimated parameters as described by Haralick. Let us consider the laser-camera

system such that the relative pose of the camera head with respect to the laser range finder be described by

$$\Theta = [{}^{\mathbf{l}}\mathbf{t}_{\mathbf{lh}}, \mathbf{\Phi}_{\mathbf{lh}}]^\top. \quad (18)$$

Here, ${}^{\mathbf{l}}\mathbf{t}_{\mathbf{lh}} = [t_x, t_y, t_z]^\top$ is a Euclidean 3-vector from l to h as expressed in frame l, and $\mathbf{\Phi}_{\mathbf{lh}} = [\theta_x, \theta_y, \theta_z]^\top$ is a 3-vector of xyz-convention roll, pitch, heading Euler angles that parametrizes the orthonormal rotation matrix ${}^l_h R$ (that rotates the frame h into frame l). The covariance of the estimated parameters Θ can thus be given as:

$$\Sigma_\Theta = \left[\frac{\partial^2 F}{\partial \Theta^2}(X,\Theta)\right]^{-1} \frac{\partial^2 F^T}{\partial X \partial \Theta}(X,\Theta)\Sigma_X \frac{\partial^2 F}{\partial X \partial \Theta}(X,\Theta) \left[\frac{\partial^2 F}{\partial \Theta^2}(X,\Theta)\right]^{-1} \quad (19)$$

Here, $X = [\mathbf{N}^{\mathbf{1}}_{\mathbf{h}}, \tilde{P}^1_l, \tilde{P}^2_l \ldots \mathbf{N}^{\mathbf{i}}_{\mathbf{h}}, \tilde{P}^1_l, \ldots]^T$ is the vector of measurements (i.e. the normals of the planes observed and the laser points lying on these planes).

3. RESULTS

We performed experiments on both simulated data as well as real data. With simulated data we can identify the optimum orientation and area of the target plane and the number of views needed to get a reasonable estimate of the extrinsic parameters.

3.1 Simulated Data

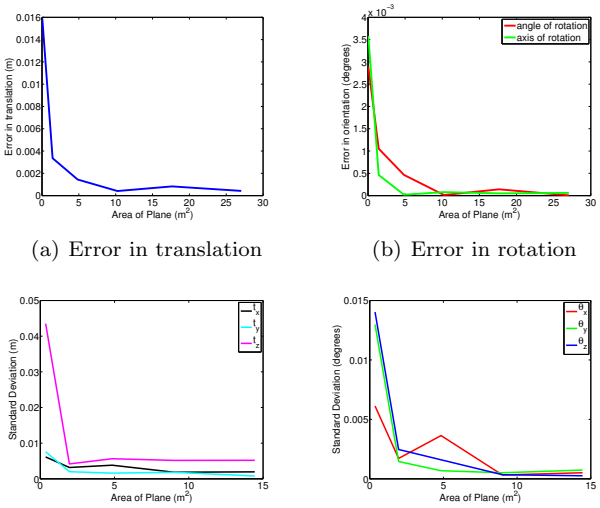
In this experiment we check the sensitivity of the algorithm to the area of the target plane and its orientation with respect to the laser-camera system and the number of normals (views) needed. We simulate a laser-camera system such that the relative pose of the camera head with respect to the laser range finder is described as:

$$\Theta = [0.5m, -1.0m, 0.8m, 85^o, 80^o, 5^o]^\top \quad (20)$$

The position, orientation and area of the plane are chosen according to the experiment to be performed. The laser points lying on the plane are computed based on the relative pose of the laser and camera head. We then add uniform gaussian noise of 10 cm to the range measurements of the laser points. These noisy points are then used to estimate the calibration parameters. The error in translation parameters is computed as the euclidean distance between the true and estimated translation vector. We use an axis angle representation for rotation parameters and the error is computed as the angle between the true and estimated axis of rotation, and the absolute difference in true and estimated angle of rotation about the axis of rotation. Following are the observations based on our simulations:

(1) Area of the plane: As shown in Fig. 6 the estimation error decreases as the area of the planar surface increases. This is because when we have a larger surface area the number of 3D laser points falling on the plane increases thereby increasing the number of constraint in the optimization (17). In practice we can stick a small (1m x 1m) checkerboard pattern on the walls available in the experimental site to get large target planes for the 3D laser data.

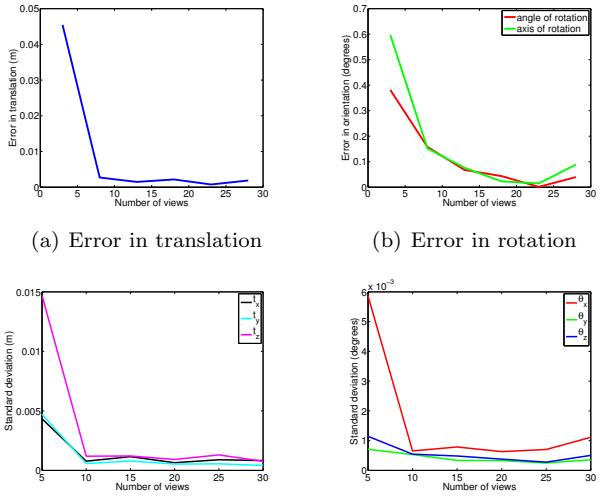
(2) Number of Normals/Views: The estimation error decreases with the increase in the number of views of the target plane (Fig. 7). Increasing the number of views increases the number of constraints in the optimization (17). Since the omnidirectional camera system is composed of six different cameras, we should take a sufficient number of planes viewed from all the cameras so that our estimate is not biased towards any one camera of the system.



(a) Error in translation (b) Error in rotation

(c) Standard Deviation of estimated translation parameters (d) Standard Deviation of estimated rotation parameters

Fig. 6. Simulation results: Error in estimation decreases as area of target plane increases. Number of views = 10



(a) Error in translation (b) Error in rotation

(c) Standard deviation of estimated translation parameters (d) Standard deviation of estimated rotation parameters

Fig. 7. Simulation results: Error in estimation decreases as number of views of target plane increases. Area of plane = $1m^2$

3.2 Real Data

The proposed extrinsic calibration method has been tested on real data collected by a vehicle mounted with a 3D laser sensor and an omnidirectional camera system, as shown in Fig. 8. We have two sets of results verifying the accuracy of the algorithm. In the first case, we considered the setup similar to the calibration setup. The calibration was performed inside a garage and checkerboard patterns were mounted on all available planar surfaces (including

side walls and ground floor). As shown in Fig. 9, the points from different planes, denoted by different colors, have been projected onto the corresponding image. In the second case we took the vehicle outside the garage and collected some data from the moving vehicle around the Ford campus. The result of projection of this 360 degree field of view point cloud over the 5 cameras of the omnidirectional camera system is shown in Fig. 10.

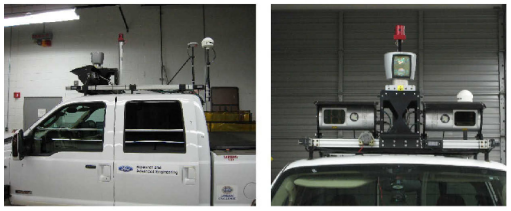


Fig. 8. Test vehicle showing sensor arrangement

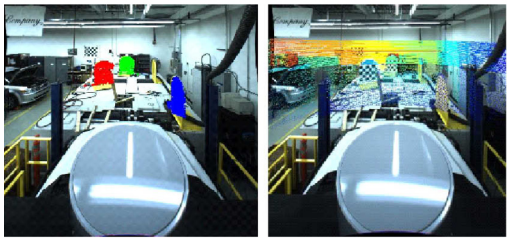


Fig. 9. Reprojection results for the calibrated system. The left panel shows the 3D points lying on the plane projected onto the images. The right panel shows the entire point cloud projected onto the image, the projected points are color coded based on depth of the point from the camera center.

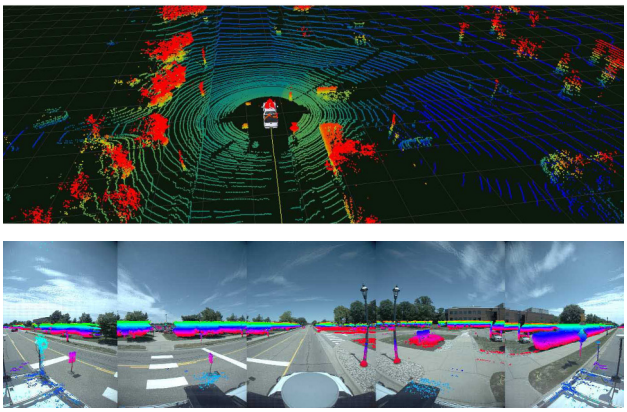


Fig. 10. The top panel is a perspective view of the Velodyne lidar range data, color-coded by height above the estimated ground plane. The bottom panel shows the above-ground-plane range data projected into the corresponding image from the Ladybug cameras.

4. CONCLUSION AND FUTURE WORKS

In this paper, we presented an extrinsic calibration method to estimate the rigid body transformation between an omnidirectional camera system and a laser scanner. The proposed method minimally requires three views of a planar pattern visible from both the camera and the laser scanner. The laser points lying on the planar surface and the normal of the plane as estimated from the image data provide a constraint on the rigid body transformation between the two sensors. Fusion of data provided by range and vision sensors constitutes an appropriate framework for mobile robot platforms to enhance various state of the art computer vision and robotics algorithms. The co-registration allows us to construct textured 3D maps of the environment, which can be used for robust navigation tasks. Moreover, the data association established here can also be used in the cost function of the state of the art ICP algorithm as an additional measure. We can also calculate the SIFT features in the image and associate these SIFT descriptors at a point in the image to the corresponding 3D point, thereby adding the appearance information to the 3D point, which will boost the various 3D object detection/classification algorithms. Thus, the method presented here is the first step to the enormous research opportunities that are available, in terms of using image and laser data together in various state of the art computer vision and robotics algorithms.

REFERENCES

Aliakbarpour, H., Nunez, P., Prado, J., Khoshhal, K., Dias, J., 2009. An efficient algorithm for extrinsic calibration between a 3d laser range finder and a stereo camera for surveillance. In: 14th International Conference on Advanced Robotics, Munich, Germany. pp. 1–6.

Fischler, M. A., Bolles, R. C., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. In: Communications of the ACM, Volume 24, Number 6.

Haralick, R. M., 1998. Propagating covariance in computer vision. In: Proceedings of the Theoretical Foundations of Computer Vision. pp. 95–114.

Levenberg, K., 1944. A method for the solution of certain problems in least squares. In: The Quarterly of Applied Mathematics 2. pp. 164–168.

Marquadrt, D., 1963. An algorithm for least-squares estimation of nonlinear parameters. In: SIAM Journal on Applied Mathematics 11. pp. 431–441.

McBride, J. R., Ivan, J. C., Rhode, D. S., Rupp, J. D., Rupp, M. Y., Higgins, J. D., Turner, D. D., Eustice, R. M., 2008. A perspective on emerging automotive safety applications, derived from lessons learned through participation in the darpa grand challenges. In: Journal of Field Robotics, 25(10). pp. 808–840.

Mei, C., Rives, P., 2006. Calibration between a central catadioptric camera and a laser range finder. In: IEEE ICRA. pp. 532–537.

Pointgrey, 2009. Spherical vision products: Ladybug3. URL `www.ptgrey.com/products/ladybug3/index.asp`

Scaramuzza, D., Harati, A., Siegwart, R., 2007. Extrinsic self calibration of a camera and a 3d laser range finder from natural scenes. In: IEEE IROS. pp. 4164–4169.

Unnikrishnan, R., Hebert, M., 2005. Fast extrinsic calibration of a laser rangefinder to a camera. Tech. rep.

Zhang, Q., 2004. Extrinsic calibration of a camera and laser range finder. In: IEEE IROS. pp. 2301–2306.

Zhang, Z., 1998. A flexible new technique for camera calibration. In: IEEE Transactions on Pattern Analysis and Machine Intelligence. pp. 1330–1334.