

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/224562418>

A Novel Multi-Planar LIDAR and Computer Vision Calibration Procedure Using 2D Patterns for Automated Navigation

Conference Paper · July 2009

DOI: 10.1109/IVS.2009.5164263 · Source: IEEE Xplore

CITATIONS

67

READS

1,756

2 authors, including:



Matthew J. Barth

University of California, Riverside

467 PUBLICATIONS 12,073 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



EV Eco-Driving [View project](#)



Digital Twin for Connected Vehicles [View project](#)

A Novel Multi-Planar LIDAR and Computer Vision Calibration Procedure Using 2D Patterns for Automated Navigation

Lili Huang, *Student Member, IEEE*, and Matthew Barth, *Senior Member, IEEE*

Abstract—In this research, we propose a unique multi-planar LIDAR and computer vision calibration algorithm. This method only requires the camera and LIDAR to observe a planar pattern at different positions and orientations. Geometric constraints of the 'views' from the LIDAR and camera images are resolved as the coordinate transformation coefficients. The proposed approach consists of two stages: solving a closed-form equation, followed by applying a non-linear algorithm based on a maximum likelihood criterion. To the author's best knowledge, this is the first paper for a multi-planar LIDAR and vision system calibration. Compared with the classical methods which use 'beam-visible' cameras or 3D LIDAR systems, this approach is easy to implement at low cost. Additionally, computer simulation and real world testing have been carried out to evaluate the performance of this approach. Lastly, application of the technique for automated navigation is presented.

I. INTRODUCTION

In recent years, numerous research projects have been carried out in real-time vehicle navigation. Mobile sensing systems consisting of a suite of sensors are used for detecting traffic conditions. These sensors provide real-time traffic surveillance measurements and play an important role in the development of Driver Assistant Systems (DAS).

One of the most common sensing techniques is the use of computer vision. Computer vision can provide a large amount of information on the surrounding environment. However, it suffers from intensity variations, narrow fields of view, and low-accuracy depth information when using stereo vision [1]. Laser range sensing (LIDAR) is another attractive technology due to its high accuracy in range measurement, its wide-area view, and low data-processing requirements [2]. LIDAR estimates distances between the on-board sensor and a variety of detected objects depending on time-of-flight of the laser light. Its measurements depend on the size and reflectivity of the target, so the probability of detection decreases with distance. Since their characteristics complement each other, it is logical to utilize both computer vision and LIDAR for detecting different objects. Sensor fusion systems are commonly used to combine the sensory data from disparate sources, so that the result will be more accurate and complete in comparison to the output of one sensor. One famous example is BOSS, the winner of the 2007 DARPA Urban Grand Challenge, which was equipped

This work is supported by the University of California Transportation Center (UCTC) Doctoral Dissertation Fellowship.

Lili Huang and Matthew Barth are with Department of Electrical Engineering, and also with the Center for Environmental Research and Technology, University of California Riverside, Riverside, CA, 92521, U.S.A. Email: lhuang@ee.ucr.edu; barth@ee.ucr.edu.

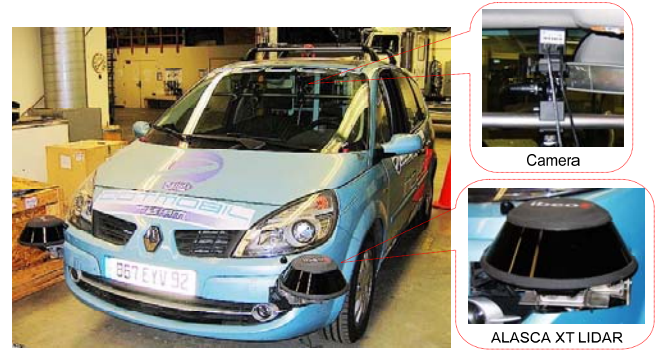


Fig. 1. The mobile sensing platform used in the experimentation. This vehicle carries one camera behind its front windshield, and two ALASCA XT LIDARs made by IBEO on the front bumper.

with GPS, long and short-range LIDAR sensors, as well as stereo cameras [3].

In order to effectively extract and integrate 3D information from both computer vision and LIDAR systems, the relative position and orientation between these two sensor modalities should be obtained. The relative geometric transformation can be solved through a calibration process [4][5]. Several approaches have been defined and utilized for LIDAR and computer vision calibration. These techniques can be roughly classified into three categories:

Visible Beam Calibration techniques are performed by observing the LIDAR beams or reflection points using cameras. The calibration system usually consists of an active LIDAR and some infrared or near-infrared cameras. A LIDAR system typically projects stripes with a known frequency, and these stripes are visible to the cameras [6-8]. This approach requires a high-cost infrared camera, which should be sensitive to the spectral emission band of the LIDAR.

Three-Dimensional LIDAR Based Calibration techniques calibrate the computer vision system with a three-dimensional LIDAR system. Various features are captured by both the camera and the LIDAR [9][10]. These features are usually in the form of corners or edges of specific calibration objects, requiring an elaborate setup. Moreover, dense LIDAR beams in both the vertical and horizontal directions are required. Therefore, these methods cannot easily be applied to single planar or sparse multi-planar LIDAR systems.

Two-Dimensional Planar-Based Calibration techniques for single planar LIDAR has previously been proposed by Zhang and Pless [17]. The calibration approach is based on

observing a plane of an object and solving distance constraints from the camera and LIDAR systems. This approach works for a single planar LIDAR only, such as the SICK 200 sensor.

To date, there does not exist any convenient calibration methods for multi-planar 'invisible-beam' LIDAR systems and computer vision systems cited in the literature. In this paper, we propose a novel calibration approach of a camera with a multi-planar LIDAR system, where the laser beams are invisible to the camera. This calibration method also works for computer vision and 3D LIDAR systems at a low cost and with high efficiency. The calibration method is then applied to a mobile sensing system with two multi-planar LIDAR systems and a single camera. The system setup is illustrated in Fig. 1.

Although there have been several calibration methods, few of them have provided a complete sensitivity analysis of the calibration procedure. As part of the calibration method proposed in this paper, we also address the effect of LIDAR noise level on the calibration process and provide an analysis of accuracy.

This paper is organized as follows: Section II examines the setup using planar planes and defines the calibration constraint. Section III describes in detail how to solve this constraint in two steps. Both a closed-form solution and a non-linear minimization solution based on maximum likelihood criterion are introduced. Experimental results with different poses are provided in Section IV. Finally, we conclude this paper with future work in Section V.

II. SENSOR ALIGNMENT

The setup for a multi-planar LIDAR and camera calibration is examined here.

A. Sensor Configuration

In the setup, an instrumented vehicle is equipped with two IBEO ALASCA XT LIDAR systems mounted on the front bumper. These LIDAR systems scan with four separate planes. The distance range is up to 200 meters, and the horizontal field of view angle of a single LIDAR is 240° [11]. The total vertical field of view for the four planes is 3.2° . The camera is mounted on the vehicle behind the front windshield.

In order to use the measurements from different kinds of sensors at various positions on the vehicle, the measurements should be transformed from their own coordinate system into some common coordinate system. In this section, we focus on obtaining the spatial relationship between video and LIDAR sensors. The geometric sensor model is shown in Fig. 2.

B. Vision, LIDAR and World Coordinate Systems

There are several coordinate systems in the overall system: the camera, the LIDAR, and the world coordinate systems.

A camera can be represented by the standard pinhole model. One 3D point in the camera coordinate denoted by

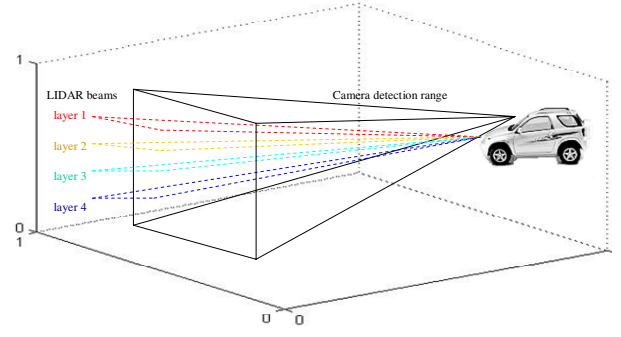


Fig. 2. Geometric model with the camera and the LIDAR.

$\mathbf{P}_c = [X_c \ Y_c \ Z_c]^T$ is projected to a pixel $\mathbf{p} = [u \ v]^T$ in the image coordinate. The pinhole model is given by [12]:

$$s\mathbf{p} \sim \mathbf{A} [\mathbf{R} \ \mathbf{t}] \mathbf{P}_c \quad \text{with } \mathbf{A} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where s is an arbitrary scale factor. \mathbf{A} is the camera intrinsic matrix defined by the coordinates of the principal point (u_0, v_0) , scale factors α and β in image u and v axes, and skewness of the two image axes γ . (\mathbf{R}, \mathbf{t}) are called extrinsic parameters. The 3×3 orthonormal rotation matrix \mathbf{R} represents the orientation of the world coordinates to the camera coordinate system. The translation matrix \mathbf{t} is a 3-vector representing the origin of the world coordinates in the camera's references. In the real world, the lens of the camera may have image distortion coefficients, which include radial and tangential distortions and are usually stored in a 5-vector [13]. In our experiments, the lens is assumed to have no significant distortion, or the distortion has already been eliminated.

The LIDAR system provides distance and direction of each scan point in the LIDAR coordinate system. Distances and directions can be converted into a 3D point denoted by $\mathbf{P}_l = [X_l \ Y_l \ Z_l]^T$ [11]. The origin of the LIDAR coordinates is the equipment itself. X , Y and Z axes are defined as forward, leftward and upward from the equipment, respectively. The camera and LIDAR reference systems are shown in Fig. 3.

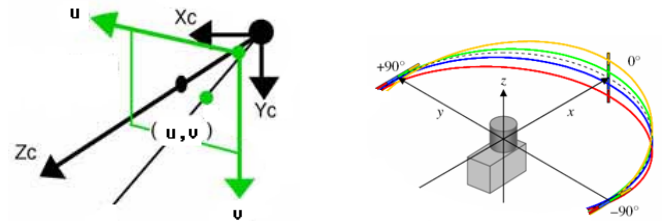


Fig. 3. (a) The camera coordinates and screen coordinate systems, and (b) the LIDAR coordinate system.

In addition to the camera and LIDAR reference systems, another coordinate system is used in calibration procedure: the world coordinate system. In the calibration process, a

checkerboard is placed in front of the sensors. The first grid on the upper-left corner of this board is defined to be the origin of the world coordinates [13].

Suppose we have a fixed point \mathbf{P} in space, which is denoted as $\mathbf{P}_c = [X_c \ Y_c \ Z_c]^T$ in the camera coordinates, and $\mathbf{P}_l = [X_l \ Y_l \ Z_l]^T$ in the LIDAR coordinates. The transformation from LIDAR coordinate to camera coordinate is given as:

$$\mathbf{P}_c = \mathbf{R}_l^c \mathbf{P}_l + \mathbf{t}_l^c \quad (2)$$

where $(\mathbf{R}_l^c, \mathbf{t}_l^c)$ are the rotation and translation parameters which relate LIDAR coordinate system to the camera coordinate system.

The purpose of our calibration work is to solve equation (2) and obtain coefficients $(\mathbf{R}_l^c, \mathbf{t}_l^c)$, so that any given points in the LIDAR reference system can be transformed to the camera coordinates.

C. Basic Geometric Interpretation

A checkerboard visible to both sensors is used for calibration. In the following sections, the planar surface defined by the checkerboard is called the *checkerboard plane*. Without loss of generality, we assume the checkerboard plane is on $Z = 0$ in the world coordinates. Let \mathbf{r}_i denotes the i th column of the rotation matrix \mathbf{R} . Then \mathbf{r}_3 is the surface normal vector of the calibration plane in the camera coordinate system [13].

Note the origin of the world coordinate is the upper-left corner of the checkerboard, and the origin of the camera coordinate is the camera itself. The translation vector \mathbf{t} represents relative position of the checkerboard's upper-left corner in the camera's reference systems. Since both \mathbf{t} and \mathbf{P}_c are points on the checkerboard plane denoted in camera coordinates, a vector \vec{v} is defined as $\vec{v} = \mathbf{P}_c - \mathbf{t}$. \vec{v} is a vector on the checkerboard plane, and \mathbf{r}_3 is orthogonal to this plane, we have:

$$\mathbf{r}_3 \cdot \vec{v} = 0 \quad (3)$$

where \cdot denotes the inner product. The geometric interpretation for equation (3) is illustrated in Fig. 4.

By taking (2) into (3), we have:

$$\mathbf{r}_3^T (\mathbf{R}_l^c \mathbf{P}_l + \mathbf{t}_l^c - \mathbf{t}) = 0 \quad (4)$$

Since point \mathbf{P}_l in LIDAR coordinates is $[X_l \ Y_l \ Z_l]^T$, from (4) we have:

$$\mathbf{r}_3^T [\mathbf{R}_l^c \ \mathbf{t}_l^c - \mathbf{t}] \begin{bmatrix} X_l \\ Y_l \\ Z_l \\ 1 \end{bmatrix} = 0 \quad (5)$$

For each LIDAR point on the checkerboard plane, equation (5) explains the geometric relationships and constraints on $(\mathbf{R}_l^c, \mathbf{t}_l^c)$. This is the basic constraints for the calibration from the LIDAR to the vision coordinate system.

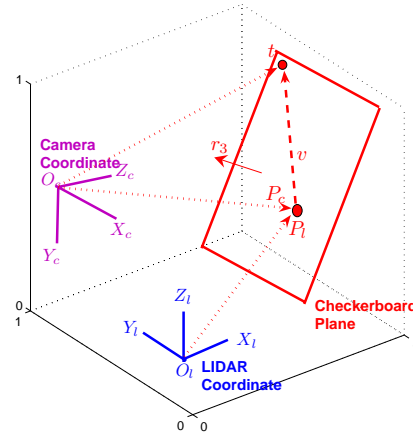


Fig. 4. Geometric interpretation of the camera coordinates, the LIDAR coordinates, and checkerboard plane.

III. CALIBRATION SOLUTION

This subsection provides the method to efficiently obtain the calibration coefficients $(\mathbf{R}_l^c, \mathbf{t}_l^c)$. We start with an analytical solution, followed by a nonlinear optimization technique based on the maximum likelihood criterion.

A. Closed-form Solution

Initially, the camera's intrinsic parameters are calibrated using a standard Camera Calibration Toolbox [13]. For each pose of the checkerboard, there's one set of camera extrinsic parameters (\mathbf{R}, \mathbf{t}) . Each (\mathbf{R}, \mathbf{t}) is determined also using the toolbox, after which \mathbf{r}_3 and \mathbf{t} in (5) are obtained.

For simplicity, let's define $\mathbf{r}_3 = [r_{31} \ r_{32} \ r_{33}]^T$, $\Delta = \mathbf{t}_l^c - \mathbf{t} = [\Delta_x \ \Delta_y \ \Delta_z]^T$, and m_{ij} be the element on the i th row, j th column in matrix \mathbf{R}_l^c . Suppose for one pose of the checkerboard, there are p LIDAR points on the checkerboard plane, denoted as $\mathbf{P}_{l,1} = [X_{l,1} \ Y_{l,1} \ Z_{l,1}]^T$, $\mathbf{P}_{l,2} = [X_{l,2} \ Y_{l,2} \ Z_{l,2}]^T, \dots, \mathbf{P}_{l,p} = [X_{l,p} \ Y_{l,p} \ Z_{l,p}]^T$. The geometric interpretation becomes an $\mathbf{Ax} = \mathbf{0}$ problem, where \mathbf{A} is a $N \times 12$ matrix, and \mathbf{x} is a 12-vector to be solved. \mathbf{A} and \mathbf{x} are given in (6).

By getting the LIDAR points $\mathbf{P}_{l,1}, \mathbf{P}_{l,2}, \dots, \mathbf{P}_{l,N}$, we can estimate \mathbf{x} using the least square method. In order to avoid the solution $\mathbf{x} = \mathbf{0}$, normalization constraints are proposed. Faugeras and Toscani [14] suggested the constraint $m_{31}^2 + m_{32}^2 + m_{33}^2 = 1$, which is singularity free. This restriction is proposed from the coincidence that $[m_{31} \ m_{32} \ m_{33}]$ is the third row of the rotation matrix \mathbf{R}_l^c . Thus solving the equation $\mathbf{Ax} = \mathbf{0}$ is transformed into minimizing the norm of \mathbf{Ax} , i.e., minimizing $|\mathbf{Ax}|$ with the restriction $m_{31}^2 + m_{32}^2 + m_{33}^2 = 1$.

$|\mathbf{Ax}|$ can be minimized using a Lagrange method [14][18]. Let $\mathbf{m}_3 = [m_{31} \ m_{32} \ m_{33}]$, and \mathbf{m}_9 be a vector containing the remaining 9 elements in \mathbf{x} . The Lagrange equation is written as:

$$L = \mathbf{A}_9 \mathbf{m}_9 + \mathbf{A}_3 \mathbf{m}_3 + \lambda (\mathbf{m}_3^T \mathbf{m}_3 - 1) \quad (7)$$

$$\mathbf{A} = \begin{bmatrix} r_{31}X_{l,1} & r_{31}Y_{l,1} & r_{31}Z_{l,1} & r_{31} & r_{32}X_{l,1} & r_{32}Y_{l,1} & r_{32}Z_{l,1} & r_{32} & r_{33}X_{l,1} & r_{33}Y_{l,1} & r_{33}Z_{l,1} & r_{33} \\ r_{31}X_{l,2} & r_{31}Y_{l,2} & r_{31}Z_{l,2} & r_{31} & r_{32}X_{l,2} & r_{32}Y_{l,2} & r_{32}Z_{l,2} & r_{32} & r_{33}X_{l,2} & r_{33}Y_{l,2} & r_{33}Z_{l,2} & r_{33} \\ \dots & & & & & & & & & & & \\ r_{31}X_{l,N} & r_{31}Y_{l,N} & r_{31}Z_{l,N} & r_{31} & r_{32}X_{l,N} & r_{32}Y_{l,N} & r_{32}Z_{l,N} & r_{32} & r_{33}X_{l,N} & r_{33}Y_{l,N} & r_{33}Z_{l,N} & r_{33} \end{bmatrix}$$

$$\mathbf{x} = [m_{11} \ m_{12} \ m_{13} \ \Delta_x \ m_{21} \ m_{22} \ m_{23} \ \Delta_y \ m_{31} \ m_{32} \ m_{33} \ \Delta_z]^T \quad (6)$$

where \mathbf{A}_3 contains the 9th to 11th columns of \mathbf{A} , and \mathbf{A}_9 contains the remaining 9 columns corresponding to \mathbf{m}_9 .

The closed-form linear solution is:

$$\lambda \mathbf{m}_3 = (\mathbf{A}_3^T \mathbf{A}_3 - \mathbf{A}_3^T \mathbf{A}_9 (\mathbf{A}_9^T \mathbf{A}_9)^{-1} \mathbf{A}_9^T \mathbf{A}_3) \mathbf{m}_3$$

$$\mathbf{m}_9 = -(\mathbf{A}_9^T \mathbf{A}_9)^{-1} \mathbf{A}_9^T \mathbf{A}_3 \mathbf{m}_3 \quad (8)$$

It's well known that \mathbf{m}_3 is the eigenvector of the symmetric positive definite matrix $\mathbf{A}_3^T \mathbf{A}_3 - \mathbf{A}_3^T \mathbf{A}_9 (\mathbf{A}_9^T \mathbf{A}_9)^{-1} \mathbf{A}_9^T \mathbf{A}_3$ associated with the smallest eigenvalue. \mathbf{m}_9 is obtained after \mathbf{m}_3 . Once \mathbf{m}_3 and \mathbf{m}_9 are known, the rotation and translation matrix $(\mathbf{R}_l^c, \mathbf{t}_l^c)$ is available.

Because of data noise, the rotation matrix \mathbf{R}_l^c may do not in general satisfy $(\mathbf{R}_l^c)^T \mathbf{R}_l^c = \mathbf{I}$. One solution is to obtain $\hat{\mathbf{R}}_l^c$, which is the best approximation of given \mathbf{R}_l^c . This $\hat{\mathbf{R}}_l^c$ has the smallest Frobenius norm of the difference $\hat{\mathbf{R}}_l^c - \mathbf{R}_l^c$, subject to $(\hat{\mathbf{R}}_l^c)^T \hat{\mathbf{R}}_l^c = \mathbf{I}$ [12].

B. Maximum Likelihood Estimation

The closed-form solution is obtained by minimizing an algebraic distance $|\mathbf{A}\mathbf{x}|$, which is not physically meaningful. In this subsection, we refine this problem through maximum likelihood function using multi-pose checkerboard planes, which is more meaningful.

In camera calibration approach, differences of image points and the corresponding projection of the ground truth point in an image are minimized [12]. This method is also valid for visible-beam LIDAR calibration [17]. In our test, the Euclidean distances from camera to the checkerboard are checked.

Note that (4) can be written as:

$$\mathbf{r}_3^T (\mathbf{R}_l^c \mathbf{P}_l + \mathbf{t}_l^c) = \mathbf{r}_3^T \mathbf{t} \quad (9)$$

where both $\mathbf{R}_l^c \mathbf{P}_l + \mathbf{t}_l^c$ and \mathbf{t} are points on the calibration plane surface, and \mathbf{r}_3 is the normal vector to this surface. Therefore, both the left and right sides of (9) are the distance between the checkerboard plane and the origin of the camera reference system.

Suppose we have n poses of the calibration plane. For the i -th pose, there is a set of $(\mathbf{r}_3, \mathbf{t})$ denoted as $(\mathbf{r}_3^i, \mathbf{t}^i)$. We assume the LIDAR points are corrupted by Gaussian distributed noise. Then the maximum likelihood function can be defined by minimizing the sum of the difference between $\mathbf{r}_3^T (\mathbf{R}_l^c \mathbf{P}_l + \mathbf{t}_l^c)$ and $\mathbf{r}_3^T \mathbf{t}$ for all the LIDAR points. Suppose for the i -th plane, there are p_i LIDAR points. Then the solution satisfies:

$$\arg \min_{\mathbf{R}_l^c, \mathbf{t}_l^c} \sum_{i=1}^n \frac{1}{p_i} \sum_{j=1}^{p_i} \left((\mathbf{r}_3^i)^T (\mathbf{R}_l^c \mathbf{P}_{l,j}^i + \mathbf{t}_l^c) - (\mathbf{r}_3^i)^T \mathbf{t}_i \right)^2 \quad (10)$$

where $\mathbf{R}_l^c \mathbf{P}_{l,j}^i + \mathbf{t}_l^c$ is the coordinate of $\mathbf{P}_{l,j}^i$ in camera reference system, according to equation (2).

By using the Rodriguez formula, the rotation matrix \mathbf{R}_l^c is transformed into a vector, which is parallel to the rotation axis and whose magnitude is equal to the rotation angle [14]. Thus $(\mathbf{R}_l^c, \mathbf{t}_l^c)$ forms a vector. Equation (10) is solved using the Levenberg-Marquardt algorithm (LMA), which provides numerical solutions to the problem of minimizing nonlinear functions [15][16]. LMA requires an initial guess for the parameters to be estimated. In our algorithm, $(\mathbf{R}_l^c, \mathbf{t}_l^c)$ in the closed form is used as this initial state. For each pose, a set of $(\mathbf{R}_l^c, \mathbf{t}_l^c)$ is obtained. The weighted average is used as an initial guess, where the scalar weight is normalized as a relative contribution of each checkerboard pose. Then LMA gives a robust solution even if the initial state starts far off the final solution.

C. Summary of Calibration Procedure

The calibration procedure proposed in this paper can be summarized as:

- 1) Place the checkerboard in view of the camera and LIDAR systems. Make sure that the plane is within the detection area of both sensors. The different poses of checkerboard cannot be parallel to each other, otherwise the parallel poses do not provide more constraints on \mathbf{R}_l^c .
- 2) Take a few measurements (images) of the checkerboard under different orientations; for each orientation, read the LIDAR points on this plane from the output.
- 3) Estimate the coefficients using the closed-form solution given in Section III A.
- 4) Refine all the coefficients using the maximum likelihood estimation in Section III B.

IV. EXPERIMENTAL RESULTS

The proposed vision-LIDAR calibration algorithm has been tested on both a computer simulation platform and with real world data.

A. Computer Simulations

The camera is assumed to have been calibrated. It is simulated to have the following property: $\alpha = 1200$, $\beta = 1000$, and the skewness coefficient $\gamma = 0$. The principal

point is (320, 240), and the image resolution is 640×480 . The calibration checkerboard consists of a 10×10 grids. The size of each square grid is $5\text{cm} \times 5\text{cm}$. The position and orientation of the LIDAR relative to the camera is also defined. The LIDAR's position in camera coordinates is $\mathbf{t}_l^c = [10 \ 150 \ 100]^T$ centimeters, and the rotation matrix \mathbf{R}_l^c is parameterized by a 3-vector rotation vector $[-85^\circ \ 10^\circ \ -80^\circ]^T$.

The LIDAR points are calculated based on the location of the camera and relative position and orientation of the checkerboard. Gaussian noise is added to the points.

Performance w.r.t. the noise level. The checkerboard plane is placed in front of camera and LIDAR. Three poses are used here. All of them has $\mathbf{t} = [-20 \ -20 \ -550]^T$ centimeter. The three rotation matrix are defined by the rotation vectors as $\mathbf{r}_1 = [170^\circ \ -5^\circ \ 85^\circ]^T$, $\mathbf{r}_2 = [170^\circ \ 15^\circ \ 85^\circ]^T$, $\mathbf{r}_3 = [170^\circ \ -25^\circ \ 85^\circ]^T$, respectively. Gaussian noise with zero mean and σ standard deviation (from 1 to 10cm) is added to the LIDAR points. The estimation results are then compared with the ground truth. For each noise level, we carried out 100 independent random trials. The averaged calibration error is shown in Fig. 5, where the calculation results are denoted as $\hat{\mathbf{R}}_l^c$ and $\hat{\mathbf{t}}_l^c$, respectively. As we can see from this figure, the calibration error increases with noise level. For $\delta < 7\text{cm}$ (which is larger than the normal standard deviation for most LIDARs), the error of norm of \mathbf{R}_l^c is less than 0.01. With three checkerboard poses, the relative translation error is less than 5% when $\sigma < 5\text{cm}$.

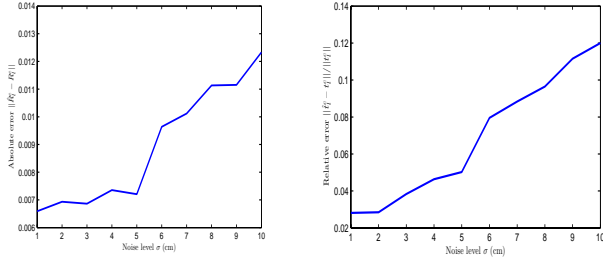


Fig. 5. Rotation and translation error w.r.t. the noise level.

Performance w.r.t. the number of checkerboard positions. The checkerboard is originally setup parallel to the image plane. Then it is rotated by 30° , where the rotation axis is randomly selected in a uniform sphere. The number of checkerboard used for calibration varies from 4 to 20. Gaussian noise with zero mean and standard deviation $\sigma = 4\text{cm}$ is added to the LIDAR points. For each number of position, 100 trials of independent rotation axis are implemented. The averaged result is illustrated in Fig. 6. This figure shows that when the number of checkerboard positions increases, the calibration error decreases.

Performance w.r.t. the orientation of the checkerboard plane. The checkerboard plane is initially set as parallel to the image plane. It is then rotated around a randomly chosen axis with angle θ . The rotation axis is randomly selected from a uniform sphere. The rotation angle θ varies from 10° to 80° , and 10 checkerboards are used for each θ . Gaussian noise

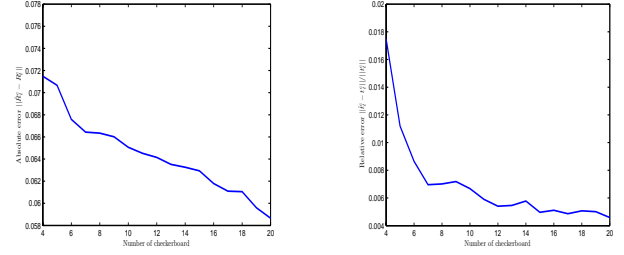


Fig. 6. Rotation and translation error w.r.t. number of checkerboard positions.

with zero mean and standard deviation $\sigma = 4\text{cm}$ is added to the LIDAR points. For each rotation angle, 100 trials are repeated and the average error is calculated. The simulation result is shown in Fig. 7. The calibration error decreases when the rotation angle increases. When the rotation angle is too small, the calibration planes are almost parallel to each other, which cause big error. When the rotation angle is too large, the calibration plane is almost perpendicular to the image plane, which makes the LIDAR measurement less precise.

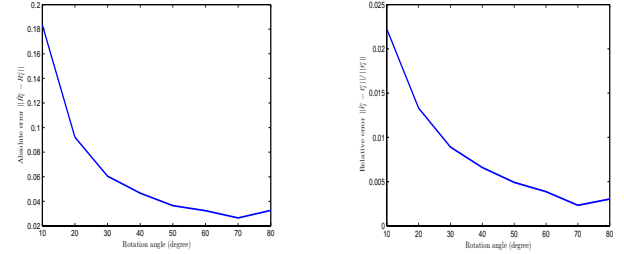


Fig. 7. Rotation and translation error w.r.t. the orientation of the checkerboard plane.

B. Real Data Calibration

The calibration method is also checked using an IBEO ALASCA XT LIDAR system and a Sony 1/3 CCD digital camera with a 6mm lens. The image resolution is 640×480 . The checkerboard plane consists of a pattern of 16×16 squares, so there are totally 256 grids on the plane. The size of each grid is $2.54\text{cm} \times 2.54\text{cm}$ ($1\text{in} \times 1\text{in}$).

20 images of the plane were taken with different orientations, and the LIDAR points are recorded simultaneously. Two examples of the calibration results are shown in Fig. 8, where the LIDAR points are mapped to image reference system using estimated \mathbf{R}_l^c and \mathbf{t}_l^c . Although the ground truth of \mathbf{R}_l^c and \mathbf{t}_l^c are not known, Fig. 8 shows that the estimation results are pretty reasonable. .

C. Application in Automated Navigation

The calibration method has been integrated into our mobile sensing system. This mobile sensing system is designed to detect and track surrounding vehicles, which is the first and fundamental step for any of the automatic traffic surveillance systems. However, object detection is a big challenge for the

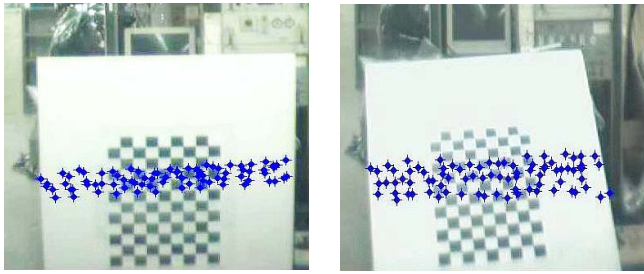


Fig. 8. Two images of the checkerboard, together with LIDAR points indicated by blue dots. The calibration method proposed in this paper is used to estimate \mathbf{R}_l^c and \mathbf{t}_l^c .

moving platform. Both the foreground and the background are rapidly changing, which makes it difficult to extract the foreground regions from the background. In our project, the sensor fusion technique is used to compensate for the spatial motion of the moving platform. Fig. 9 gives two images from the vehicle detection video.

In Fig. 9(a), there are totally four vehicles detected by the LIDAR, where the farthest vehicle is 55 meters away from our mobile sensing system. It's hard to obtain the vehicle's distance and orientation from an image only. The LIDAR points provide a reliable estimation of this vehicle's position. In Fig. 9(b), one car parallel to the probe vehicle is detected by the LIDAR. Meanwhile, it is partially visible in the image, together with its shadow. Although this vehicle is hardly recognizable in the image, with a wide angle of view, the LIDAR data provide enough information to reconstruct the location of the vehicle.



Fig. 9. Two images of sensor fusion in automatic navigation. The red rectangle is an enlarged image of the detected area.

Our experiment results illustrate that the calibration algorithm is successful. For the automatic navigation application, the LIDAR provides reliable distance information. The sensor fusion system combining LIDAR and computer vision information sources presents distance and orientation information. This system is helpful for vehicle detection and tracking applications. More details about the application in our automated navigation system are introduced in [19].

V. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a novel calibration algorithm to obtain the geometry transformation between a multi-plane LIDAR system and the camera vision system. This calibration method requires the LIDAR and camera to observe a checkerboard simultaneously. A few checkerboard

poses are observed and recorded. The proposed calibration approach has two stages: closed-form solution followed by the maximum likelihood criterion based optimization. Both simulation and real world experiments have been carried out. The experiment results show that the calibration approach is reliable. This method can be applied for both multi-plane and 3D LIDAR sensor fusion systems.

Future work will concentrate on developing the mobile sensing system. Since LIDAR provides the dimension and orientation information, and camera image records colors and textures, these two sensor data will be combined for vehicle recognition and activity analysis. The occlusion problem will also be considered as future research.

REFERENCES

- [1] M. Y. Kim, H. Cho and H. Lee, "An Active Trinocular Vision System for Sensing Mobile Robot Navigation Environments," *IEEE Intelligent Robots and Systems Proceedings*, pp. 1698-1703, 2004.
- [2] Young-Kee Jung and Yo-Sung Ho, "Traffic Parameter Extraction Using Video-based Vehicle Tracking", *IEEE/IEEE/ISAI International Conference on Intelligent Transportation Systems*, pp. 764-769, 1999.
- [3] W. R. Whittaker, "Red Team DARPA Grand Challenge 2005 Technical Paper," Carnegie Mellon University, *CSC Technical Report*, August 2005.
- [4] H. Baltzakis, A. Argyros, and P. Trahanias, "Fusion of Laser and Visual Data for Robot Motion Planning and Collision Avoidance," *Machine Vision and Application*, pp. 431-441, 2003.
- [5] Y. Liu and R. Emery, "Using EM to Learn 3D Environment Models with Mobile Robots," *Proc. 18th International Conference on Machine Learning*, 2001.
- [6] Olli Jokinen, "Self-Calibration of a Light Striping System by matching Multiple 3-D Profile Maps," *Proc. Second International Conference on 3-D Digital Imaging and Modeling*, pp. 180-190, 1999.
- [7] Mitsuhiro Hayashibe and Yoshihiko Nakamura, "Laser-Pointing Endoscope System for Intra-Operative 3D Geometric Registration," *Proc. International Conference on Robotics and Automation*, pp. 1543-1548, May 2001.
- [8] Mirko Mahlich, Roland Schweiger, Werner Ritter and Klaus Dietmayer, "Sensorfusion Using Spatio-Temporal Aligned Video and Lidar for Improved Vehicle Detection," *Proc. of Intelligent Vehicles Symposium*, pp. 424-429, Jun. 2006.
- [9] S. Wasielewski and O. Strauss, "Calibration of a Multi-sensor System Laser Rangefinder / Camera," *Proc. of Intelligent Vehicles Symposium*, pp. 472-477, 1995.
- [10] R. Unnikrishnan and M. Hebert, "Fast Extrinsic Calibration of a Laser Rangefinder to a Camera," *tech. report CMU-RI-TR-05-09*, Robotics Institute, Carnegie Mellon University, July, 2005.
- [11] The Laser Scanner Product Overview, <http://www.ibeo-as.com/english/products.asp>.
- [12] Zhengyou Zhang, "A Flexible New Technique for Camera Calibration," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, pp. 1330-1334, 2000.
- [13] Jean-Yves Bouguet, "Camera Calibration Toolbox for Matlab," 2003.
- [14] Faugeras, O.D. and Toscani, G., "Camera Calibration for 3D computer vision," *Proc. International Workshop on Industrial Applications of Machine Vision and Machine Intelligence*, pp. 240-247, 1987.
- [15] K. Levenberg, "A method for the solution of certain problems in least squares," *Quarterly Applied Math*, pp. 164-168, 1944.
- [16] J. More, "The Levenberg-Marquardt Algorithm, Implementation and Theory," In G.A. Watson, editor, *Numerical Analysis, Lecture Notes in Mathematics 630*, Springer-Verlag, 1977.
- [17] Qilong Zhang and Robert Pless, "Extrinsic Calibration of a Camera and Laser Range Finder (improves camera calibration)," *IEEE Proceedings of International Conference on Intelligent Robots and Systems*, pp. 2301-2306, Sep. 2004.
- [18] Carlos Ricolfe Viala and Antonio Jose Sanchez Salmeron, "Performance Evaluation of Linear Camera Calibration Techniques," *IEEE Proceedings on World Automation Congress*, Vol. 18, pp. 49-54, 2004.
- [19] Lili Huang and Matthew Barth, "Tightly-Coupled LIDAR and Computer Vision Integration for Vehicle Detection," *IEEE Intelligent Vehicle Symposium*, 2009.