A Simulation Study on Calibration of A LiDAR with Respect to A Camera by Using Point and Plane Constraints

Fumio Itami $^{1}$ and Takaharu Yamazaki $^{2}$

$^{1}$Saitama Institute of Technology
$^{2}$Affiliation not available

December 7, 2023

Abstract

This letter provides a simulation study on calibration of a 3D LiDAR with respect to a camera, with point-topoint and point-on-plane correspondences for calibration constraints, by extending our previously proposed 2D LiDAR calibration methods. The calibration performance is examined in terms of vertical scan view and resolution, especially in narrow/sparse scan space, in the presence of sensor noise. It contributes to a guidance on a use of various sensors and dataset depending on contexts and applications with different scan view and resolution, and an understanding of the effect of the different constraints on calibration. As a result, the combination of the point and plane constraints yields better performance than either of them.
A Simulation Study on Calibration of A LiDAR with Respect to A Camera by Using Point and Plane Constraints

Fumio Itami¹* and Takaharu Yamazaki¹
¹Department of Information Systems, Saitama Institute of Technology, 1690, Fusaiji, Fukaya, Saitama, 369-0293, Japan
*Member, IEEE

Abstract—This letter provides a simulation study on calibration of a 3D LiDAR with respect to a camera, with point-to-point and point-on-plane correspondences for calibration constraints, by extending our previously proposed 2D LiDAR calibration methods. The calibration performance is examined in terms of vertical scan view and resolution, especially in narrow/sparse scan space, in the presence of sensor noise. It contributes to a guidance on a use of various sensors and dataset depending on contexts and applications with different scan view and resolution, and an understanding of the effect of the different constraints on calibration. As a result, the combination of the point and plane constraints yields better performance than either of them.

Index Terms—Camera, LiDAR, Calibration.

I. INTRODUCTION

In computer vision and robotics, a number of sensor calibration methods have been proposed. In particular, the calibration of a Light Detection And Ranging sensor (LiDAR) which obtains depth information in surroundings, with respect to a camera which obtains vision information in surroundings, is important to understand surroundings with their complimentary integration. So far, the calibration methods which use calibration targets have been given as follows. One of the most applicable targets is a checkerboard, which, in most cases, is identical to one used in the camera calibration. The calibration methods [1]-[4] use a checkerboard target, for the 2D LiDAR calibration. On the other hand, the calibration methods [5]-[8] use other calibration targets for the 2D LiDAR calibration. Moreover, a checkerboard or a similar planar board is utilized in [9]-[12], and other calibration targets are utilized in [13]-[16], for the 3D LiDAR calibration. In addition, instead of using a usually expensive 3D LiDAR, a use of a "rotating" 2D LiDAR [17][18] is another selection for constructing a 3D perception system.

In the above calibration methods, the authors have also proposed a calibration method which uses a point-like target to obtain point-to-point correspondences [8], and another calibration method which uses a checkerboard target to obtain point-on-plane correspondences [4], for more sufficient constraints than traditional ones, in the 2D LiDAR calibration. However, in these methods [4][8], and the other traditional methods too, the relation between sensor-calibration performance and various sensor-specific dataset, such as those with 2D/3D scan space, narrow/broad scan view, sparse/dense scan resolution, etc., especially in the vertical scan, has not sufficiently been discussed. The examination of the relation contributes to a guidance on a use of various sensors and dataset, including a rotating 2D LiDAR or a 3D LiDAR with such mechanicals, depending on contexts and applications with different scan view and resolution. Furthermore, the methods [4][8] have the constraints different from each other, i.e., point or plane. Thus, the examination of their performance also contributes to an understanding of the effect of the different constraints on calibration, which has not sufficiently been discussed either, even in the recent study [9]-[12], where point-on-plane and point-to-point constraints are utilized, but limited to on a planar board.

Therefore, this letter provides a simulation study on the previously proposed 2D LiDAR calibration methods [4][8] with an extension, where point-to-point and point-on-plane correspondences for calibration constraints are obtained in more general 3D scan space, i.e., for a use of a 3D LiDAR. We specifically examine the calibration performance in terms of the vertical scan view and resolution, especially in narrow/sparse scan space, in the presence of sensor noise, but not other human/mechanical-specific noise which can be caused by human manipulation and the mechanical precision in a calibration or sensing system [4][8], to examine the essential relation between the calibration performance and the dataset, i.e., the best or limit of the performance with the dataset.

II. SENSOR CALIBRATION METHODOLOGY

A. Calibration of a camera

The calibration of a camera [19] is to estimate the internal camera characteristics. Usually a checkerboard is utilized as a calibration target, to obtain point-to-point correspondences in the world frame and in the image frame (colored green), as shown in Fig. 1. A point \( P = [x, y, z]^T \) in the world frame, which is usually an intersection point on the checkerboard pattern, is transformed to a point \( P_c = [x_c, y_c, z_c]^T \) in the camera frame, by using the rotation and translation parameters, i.e., the camera extrinsic parameters. A point in the camera frame is then transformed to a point \( p = [u, v]^T \) in the image frame, i.e., an image point, by using the camera intrinsic parameters.

Specifically, the camera intrinsic parameters, such as the focal length, the principal point, and the skewness, constitute the matrix \( K \) which transforms \( P_c \) to \( p \). The distortion parameters are also incorporated when the camera lens distortion is present. The camera extrinsic parameters, which depend on a checkerboard pose, constitute the rotation matrix \( R_c \) and the translation vector \( t_i \) in the i-th pose, which transform \( P \) to \( P_c \). These parameters are estimated by using the point-to-point constraints in the camera calibration, prior to the LiDAR calibration.
B. Calibration of a LiDAR

In the calibration of a LiDAR with respect to a camera, a point \( P_2 = [x_t, y_t, z_t]^T \) in the LiDAR frame, i.e., a laser point, is transformed to a point \( P'_2 = [x_c, y_c, z_c]^T \) in the camera frame, by using the rotation matrix \( R_{kc} \) and translation vector \( t_{kc} \), i.e., the LiDAR extrinsic parameters, which are estimated in the following.

Our first method for the LiDAR calibration [8], referred to as the "point" method in this letter, utilizes the pre-estimated camera intrinsic parameters, and the point-to-point correspondences in the LiDAR and camera frames, with the same rotation matrix \( R \) and translation vector \( t \) in the "plane" method in this letter, utilizes the pre-estimated camera intrinsic parameters, and the point-on-plane correspondences in the LiDAR and camera frames, (colored red), as shown in Fig. 1. Mathematically, the LiDAR extrinsic parameters are estimated by minimizing the distance between the projected image points, calculated from the measured laser points, and the detected image points as

\[
\arg\min_{R_{kc}, t_{kc}} \sum_{i=1}^{N_p} \left\| [u^k_i, v^k_i]^T - [u^m_i, v^m_i]^T \right\|^2,
\]

where \([u^k_i, v^k_i] \) is a projected image point calculated from a measured laser point \( P'_k \), \([u^m_i, v^m_i] \) is a detected image point, and \( N_p \) denotes the number of target poses. In practice, the laser points and the image points are obtained by moving a "point-like" target onto a scan plane of a LiDAR, which can lead to spatio-temporal mismatch between a laser point and an image point [8].

Our second method for the LiDAR calibration [4], referred to as the "plane" method in this letter, utilizes the pre-estimated camera rotation matrix \( R \) and translation vector \( t \), i.e., the camera extrinsic parameters in the \( i \)-th target pose. Furthermore, the point-on-plane correspondences in the LiDAR and camera frames, with the same checkerboard target as in the camera calibration, are also utilized (colored blue), as shown in Fig. 1. Mathematically, the LiDAR extrinsic parameters are estimated as

\[
\arg\min_{R_{kc}, t_{kc}} \sum_{i=1}^{2N_p} \sum_{j=1}^{N_p} \frac{1}{N_p} \sum_{k=1}^{N_p} \left\| r_i^T (R_{kc} P_i^{j,k} + t_{kc}) - r_i^T t_j \right\|^2,
\]

where the vector \( r_i \) is obtained from the rotation matrix \( R \), and the number of target poses is denoted as \( N_p \). The \( j \)-th laser point on the target in the \( i \)-th pose is denoted as \( P_i^{j,k} \) in the LiDAR frame, and the number of laser points on the target in the \( i \)-th pose is denoted as \( N_p \).

Note that, in the plane method, the laser points obtained in the vertically rotated LiDAR frame, i.e., vertical laser points, are also utilized in (2), where a vertical laser point, relative to the original LiDAR frame, is denoted as \( P_i^{j,k} \), and the \( (i - N_p) \)-th pose of the target, with \( r_i = r_i^{-(i-N_p)} \), and \( t = t^{-(i-N_p)} \), and \( i = N_p + 1, \ldots, 2N_p \). Also, the plane method is particularly referred to as "baseline" method in this letter, when laser points only in the originally positioned LiDAR frame are utilized with the number \( N_p \) of target poses. In the plane method, the rotation/translation discrepancy, between the originally positioned and vertically rotated LiDAR frame, can be caused in practice, depending on the mechanical precision [4].

III. COMPUTER SIMULATION

In this simulation, the calibration performance of the point, plane, baseline (part of plane) methods, mentioned above, and the combination of them, is examined in detail for a use of a 3D LiDAR, where, firstly, the vertical scan view varies whereas the vertical scan resolution is fixed, and secondly, the vertical scan resolution varies whereas the vertical scan view is fixed.

A. Ground truth and dataset

The camera intrinsic parameters are set as follows, similarly to [4][8]. The focal length \( \alpha \), \( \beta \) in pixel is set to 500, and the principal point \( u_0, v_0 \) in pixel is 320, 240. The radial distortion coefficients \( k_1, k_2 \) are -0.3, 0.05, the skewness \( i \) of the axes is set to 0, and the tangential distortion is also ignored. The camera extrinsic parameters, the rotation and translation of the checkerboard relative to the camera frame, are randomly generated in each checkerboard pose. The LiDAR extrinsic parameters are also set as in [4][8], in the following. The rotation ones \( r_x, r_y, r_z \) (degree) around each axis are set to 10, 0, and the translation ones \( t_x, t_y, t_z \) (mm) in each axis are set to 0, 250, 100, which constitute the rotation matrix and translation vector, respectively.

By using the above ground truth, the following dataset is generated. As for the camera calibration, the image points, which are the projection of the points in the world frame, are generated in each checkerboard pose. The checkerboard has the block size of 60 mm and the intersection points of 10x7 in the world frame, and is positioned within a few meters from the camera, similarly to [4][8]. As for the LiDAR calibration, in the point method, laser points are randomly generated, and then corresponding image points are generated, with a point target positioned in the same manner as the checkerboard. In the plane method, the laser points, not only in the original LiDAR frame but also in the vertically rotated LiDAR frame, are generated in each of the same checkerboard poses as in the camera calibration. Furthermore, Gaussian noise with zero mean is added to the image and laser points in both methods, as in [4][8].

B. Parameter estimation

The estimation of the camera and LiDAR parameters was conducted by using the above dataset. First, the estimation of the camera intrinsic and extrinsic parameters was carried out, and then the estimation of the LiDAR extrinsic parameters was carried out. The estimation problem was solved by using the Levenberg-Marquardt Algorithm.
Fig. 2. Box-whisker plot of the distributions of the estimation error of the obtained LiDAR extrinsic rotation and translation parameters: The three sets of the five consecutive distributions are shown, from left to right, in the vertical scan view of ±2, ±4, and ±10 degrees with the fixed scan resolution of 2 degrees, respectively, in the fixed horizontal scan view of ±10 degrees with the fixed scan resolution of 0.5 degrees.

[20][21], and the estimation of the parameters was conducted by 100 trials where the estimation error of the obtained LiDAR extrinsic parameters in rotation and translation was calculated as the norm of the vector of 3 elements which were the differences between the ground truth and the estimated parameters, as in [4][8].

The box-whisker plots of the distributions of the estimation error are shown in Fig. 2 and Fig. 3, where those of the rotation and translation error are shown at the top and bottom, respectively. The five consecutive distributions, from left to right, are the baseline method, the plane method, the point method, and moreover, the combination of the baseline and point methods, and, the combination of the plane and point methods, respectively. The standard deviation of the image noise is fixed to 0.5 pixels, and that of the LiDAR measurement noise is 30 mm, similarly to [4][8]. The number of checkerboard poses is 20 in the baseline, plane, and point methods, and the number of image/laser (target) points in one planar scan is also 20 in the point method, whereas the number of laser points on the checkerboard, in one planar scan, depends on the horizontal scan view and resolution in the plane and baseline methods, which is 40 from the following simulation setting.

In Fig. 2, the three sets of those (five) distributions are shown, from left to right, in the vertical scan view of ±2, ±4, and ±10 degrees with the fixed scan resolution of 2 degrees, respectively, in the fixed horizontal scan view of ±10 degrees with the fixed scan resolution of 0.5 degrees.

Fig. 3. Box-whisker plot of the distributions of the estimation error of the obtained LiDAR extrinsic rotation and translation parameters: The three sets of the five consecutive distributions are shown, from left to right, in the vertical scan resolution of 10, 5, and 2 degrees with the fixed scan view of ±10 degrees, respectively, in the same horizontal scan view and resolution as in Fig. 2. Note that the right set of those distributions in Fig. 2 is identical to the right set in Fig. 3.

C. Discussion

Firstly, it is observed in Fig. 2 where the vertical scan view varies, that when the vertical scan view is ±2 degrees, i.e., it is almost a planar 2D scan, the estimation error in the baseline method is very large and even that in the plane method is large, compared to the point method. On the other hand, when the vertical scan view is ±10 degrees, the estimation error in the baseline and plane methods is almost comparable to that in the point method. Note also that the estimation error in the point method gradually decreases, as the vertical scan view and thus the total number of laser points increases. It implies that the constraints in the baseline method are very weak and those in the plane method work well for their complement, in a 2D-like scan. On the other hand, the constraints in the point method are comparably strong in a 2D-like scan, and those in a general 3D scan do not have much effect, in reduction, on the estimation error.
Our future work.

etc. We leave a development of such a real and precise system for

hand, the human/mechanical-specific noise, which can be caused by

etc., as well as sensor specifications or mechanicals. On the other

such as, a small/large calibration target, narrow/broad sensing space,

and dataset, depending on their scenarios and applications, because

various factors, such as, a small/large calibration target, narrow/broad sensing space,

etc., as well as sensor specifications or mechanicals. On the other

hand, the human/mechanical-specific noise, which can be caused by

human manipulation and the mechanical precision in a calibration or

sensing system, was omitted. This type of noise can be suppressed by

automating or mechanizing the system, such as, utilizing or designing

robot arms to move a calibration target, infrared cameras to visualize

laser beam, rotating mechanicals to set a direction of laser beam,

etc. We leave a development of such a real and precise system for

our future work.

IV. CONCLUSIONS

This letter provided a simulation study on calibration of a 3D

LiDAR with respect to a camera, with point-to-point and point-on-

plane correspondences for calibration constraints, by extending our

previously proposed 2D LiDAR calibration methods. Consequently,

the results implied that, firstly, the point constraints were effective

than the plane constraints in a 2D-like scan, i.e., vertically narrow

scan view, and secondly, the vertical scan view had more effect

than the vertical scan resolution, and thirdly, the point and plane

constraints differed from each other, so that the combination of them

was effective on performance improvement.

It is expected that the results help users effectively utilize sensors

and dataset, depending on their scenarios and applications, because

the scan view/resolution can practically be affected by various factors,

such as, a small/large calibration target, narrow/broad sensing space,

etc., as well as sensor specifications or mechanicals. On the other

hand, the human/mechanical-specific noise, which can be caused by

human manipulation and the mechanical precision in a calibration or

sensing system, was omitted. This type of noise can be suppressed by

automating or mechanizing the system, such as, utilizing or designing

robot arms to move a calibration target, infrared cameras to visualize

laser beam, rotating mechanicals to set a direction of laser beam,

etc. We leave a development of such a real and precise system for

our future work.

REFERENCES


Finder (improves camera calibration),” Proc. IEEE International Conference on


Calibration of a Camera and a Laser-Rangefinder,” IEEE Transactions on Pattern