The 2021 World Congress on
Advances in Nano, Bio, Robotics and Energy (ANBRE21)
GECE, Seoul, Korea, August 23-26, 2021

Extrinsic Calibration of LiDAR and Camera using Multiple Traffic Signs

*Wonho Song[1)], Changki Sung[2)], Euigon Jung[1)], Minho Oh[1)] and Hyun Myung[3)]

[1), 3)] *School of Electrical Engineering, KAIST, Daejeon 34141, Rep. Korea*
[2)] *Robotics Program, KAIST, Daejeon 34141, Rep. Korea*
[3)] *hmyung@kaist.ac.kr*

ABSTRACT

For fusing accurate multimodal measurements, an external calibration of various sensors is required. Extrinsic transformations, on the other hand, can drift slowly when driving or maintaining, reducing the precision of the sensory system. We propose an extrinsic calibration tool for a camera and a LiDAR sensor using multiple traffic signs on a moving vehicle. We demonstrate that traffic signs can be detected accurately using a camera and a LiDAR, thus contributing to the extrinsic calibration. By using a normal vector of segmented signs from the image and point cloud, we are able to estimate an accurate extrinsic matrix of two sensors. To demonstrate the result of our method, we use our own experimental dataset. This work can provide the essential performance of using existing known objects on the street in order to get a precise measurement of extrinsic calibration of LiDAR and camera sensors.

1. INTRODUCTION

By providing supplementary information, a camera and a LIDAR assist any kind of robot, including the autonomous vehicle, in perceiving their surroundings. Because cameras lack the depth information that LIDARs provide and LIDARs lack the color, texture, and appearance information that cameras provide, systems that combine the two sensing modalities can compensate for each other's shortcomings by using their respective advantages. Perception, multi-sensor state estimation, mapping, and localization, including SLAM are all possible with a camera and a LIDAR working together. For combining information from each of these sense modalities, estimating the extrinsic calibration of two different sensors is crucial.

There are two types of methods, target-based and target-less, for extrinsic calibration of LiDAR and camera. In target-based approaches, special calibration targets having a specified size, such as polygonal planar boards [1], boxes [2], checkerboard

[1) 2)] Graduate Student
[3)] Professor

patterns [3], or a simple printed circle [4] are used. Target-less techniques [16], [17], on the other hand, extract features for correspondence calculation directly observing the natural environment, eliminating the need for a calibration object. Natural environmental features are used as an input to the calibration operation in target-less approaches. However, the accuracy of the target-less method is commonly lower than the target-based method because detecting the features from the natural environment is reasonably hard. Even a well-calibrated system needs some recalibration due to vibration on the roads and some sensor artifacts, so to avoid complex re-calibration offline process we propose an automatic, online registration method which is able to precisely calibrate LIDAR and camera sensors on any urban environment.

In this paper, we propose a novel target-less and fully automatic extrinsic calibration method between a camera and a LiDAR mounted on a moving car. Our technique requires multiple traffic signs in an urban environment, meanwhile the method calculates all the necessary registration parameters online. Section 2 presents our proposed method to calibrate a camera and a LiDAR sensor mounted on a self-driving vehicle. Section 3 discusses the evaluation of the proposed calibration method. Section 4 contains the conclusion and alludes to future work.

2. THE PROPOSED APPROACH

In this section, we describe details of the proposed method step by step. The overview of the proposed calibration method for a camera-LiDAR sensor unit is shown in Fig. 1. First, 2D images and 3D point cloud data of a traffic sign are captured from a camera and a LiDAR sensor module. We acquire multiple traffic sign data at different poses. Then, the traffic signs are detected and segmented in both image and point cloud. Finally, estimate extrinsic calibration matrix by using a normal vector of plane and corners of the segmented result.

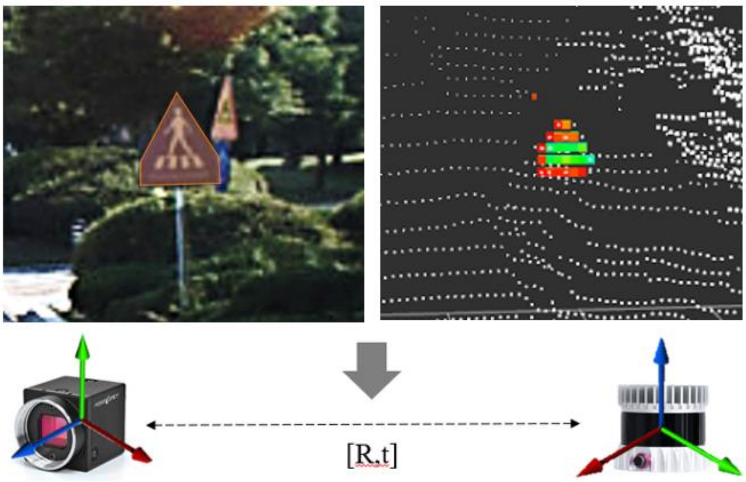

Fig. 1 Overview of the proposed calibration method

2.1 Extracting traffic signs in a camera image

There is a trained YOLO [7] that can recognize and classify 80 COCO objects.

However, the original YOLO architecture has been modified and trained to detect traffic signs for this project. We used YOLOv3 and the German Traffic Sign Recognition Benchmark (GRSRB) dataset. The GTSRB dataset [8] is used for training (80%) and validation (20%) of the neural network. The bounding box predicts the label and confidence score of each of the traffic signs detected in those boxes. We simply used canny edge detection and color threshold to segment detected results. The results are shown in Fig. 2.

Fig. 2 Traffic sign detection and segmentation results

2.2 Extracting traffic signs in LiDAR point cloud

To extract the sign point cloud data candidate group, data accumulation, and voxelization, intensity filtering, and Euclidean clustering are in order. The existing sign extraction algorithm using the lidar sensor extracted only the signs parallel to the vehicle. However, since there are many signs installed not parallel to the vehicle in an urban environment, LiDAR data is accumulated and used as a method to improve this point.

First, accumulation and voxelization of lidar data are performed. At this time, since the data of the lidar sensor acquires geographical information based on the lidar, it is converted into global coordinates and accumulated. In this way, when lidar data is accumulated for a predetermined threshold, voxelization is performed through a predetermined size voxel.

Signs have the characteristic of having a significantly higher reflectance compared to general objects on the road. Therefore, it is very effective as a method of extracting point cloud data that can correspond to a sign through intensity filtering. As the intensity of LiDAR data changes according to the surrounding environment, the intensity value in LiDAR data is in the top 0.15% so that it can be used according to changing situations, rather than setting a specific value in the filtering process to extract the point cloud data. Then, the point clusters are clustered using the Euclidean clustering method.

Finally, in order to make point cloud to place in same plane, we implemented a RANSAC [9] plane fitting algorithm to further refine the detection. Then projecting the noise point cloud onto the ideal plane as shown in Fig. 3.

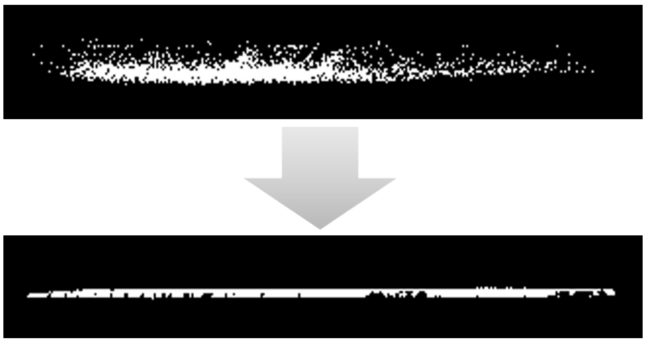

Fig. 3 Point cloud projection on a plane estimated with RANSAC

2.3 Extrinsic calibration matrix estimation

We used a similar technique mentioned in [10] in order to calculate the translation and rotation of two sensors. The normal vectors of all plane pairs are aligned to determine the relative orientation between the camera and the LiDAR. We can find N sets of 3D traffic sign corners because we have all N views of sign images. All reprojected points are reconstructed with regard to the camera coordinate system after the corner points are reprojected into 3D space. Then using a normal vector of each plane and 3D corner point of point cloud data, rotation and translation between two sensors are estimated. By using the estimated calibration matrix, the point cloud projection result is shown in Fig. 4.

Fig. 4 LiDAR point cloud projected on image plane

3. EXPERIMENT RESULT

We used our own sensor setup to test the proposed method. Ouster LiDAR and a pointgrey camera are mounted on the top of an autonomous vehicle. The sensor setup is shown in Fig. 5.

Fig. 5 Experiment sensor system

We compared our method with the publicly available calibration method [11] for lidar to stereo camera calibration. The results were evaluated using colorized accumulated point cloud as shown in Fig. 6. Known checkerboards with different locations were able to use to capture the error. The compared existing open-source method was just using a single board. We made the board tilted and placed in the center in order to get the best result using the existing method. Three left-side traffic signs and two right-side traffic signs were used for the result of our proposed method. The result is shown in Table 1.

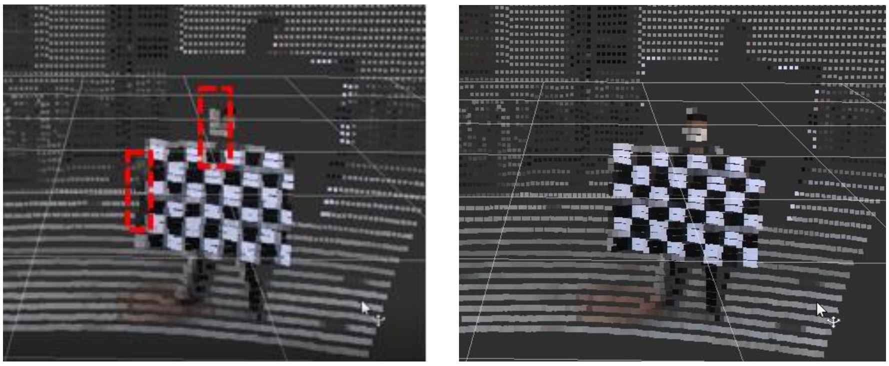

Fig. 6 Colorization of the point cloud using calibration result. Left shows the miscalibrated result using the existing method (errors: red box), and right shows the calibration result from the proposed method

Table 1 Result of the proposed method compare to existing method

Extrinsic calibration method	**RMSE [m]**		
	x	y	z
Guindel et al. [11]	0.0243	0.0287	0.0309
Proposed method	0.0194	0.0242	0.0193

4. CONCLUSIONS

We presented an extrinsic calibration tool for a camera and a LiDAR, and proposed three configurations to estimate the sensor poses from simultaneous detections of multiple calibration board locations. Experiments on a setup with two sensing modalities show that any configurations can provide good calibration results when data are collected in an urban environment with multiple traffic signs. Furthermore, the results with five different traffic sign locations show that the expected RMSE of translation is approximately 2 cm. RMSE of y were similar compare to the existing method because all traffic signs were located on the side of the road. The future work involves estimating results using open-dataset such as KITTI dataset, and extracting and matching more features, including poles and lane markings.

REFERENCES

[1] Y. Park, S. M. Yun, C. S. Won, K. Cho, K. Um, and S. Sim, “Calibration between color camera and 3d lidar instruments with a polygonal planar board,” *Sensors*, vol. 14, no. 3, pp. 5333-5353, 2014.
[2] Z. Pusztai, I. Eichhardt, and L. Hajder, “Accurate calibration of multi-lidar-multi-camera systems,” *Sensors*, vol. 18, no. 7, pp. 2139, 2018.
[3] A. Geiger, F. Moosmann, O. Car, and B. Schuster, “Automatic camera and range sensor calibration using a single shot,” in *IEEE International Conference on Robotics and Automation*, 2012, pp. 3936–3943.
[4] A. H., B. L.D., and B. B, “Automatic calibration of a range sensor and camera system,” *in Second International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission*, October 2012, pp. 286292.
[5] G. Pandey, J. R. McBride, S. Savarese, and R. M. Eustice, “Automatic extrinsic calibration of vision and lidar by maximizing mutual information,” *Journal of Field Robotics,* vol. 32, no. 5, pp. 696–722, 2015.
[6] N. Schneider, F. Piewak, C. Stiller, and U. Franke, “RegNet: Multimodal sensor registration using deep neural networks,” *in IEEE Intelligent Vehicles Symposium (IV)*, 2017, pp. 1803–1810.
[7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” *in CVPR*, 2016, pp. 779-788.
[8] J.Stallkampa, M.Schlipsinga, J.Salmena, and C.Igelb, “Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition,” *Neural Networks*, vol. 32, pp. 323-332, 2012
[9] Martin A Fischler and Robert C Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
[10] X. Gong, Y. Lin, and J. Liu, ‘‘3D LIDAR-camera extrinsic calibration using an arbitrary trihedron,’’ *Sensors*, vol. 13, no. 2, pp. 1902–1918, Feb. 2013.
[11] C. Guindel, J. Beltran, D. Martin, and F. Garcia, “Automatic extrinsic calibration for lidar-stereo vehicle sensor setups,” *in IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017, pp. 1–6.