2D LiDAR and Camera Fusion in 3D Modeling of Indoor Environment

Juan Li, Xiang He, Jia Li
Department of Electrical and Computer Engineering
Oakland University
Rochester, MI 48309, U.S.A
jli234567@oakland.edu, xhe2@oakland.edu, li4@oakland.edu

***Abstract*—Detailed 3D modeling of indoor scene has become an important topic in many research fields. It can provide extensive information about the environment and boost various location based services, such as interactive gaming and indoor navigation. This paper presents an indoor scene construction approach using 2D line-scan LiDAR and entry-level digital camera. Both devices are mounted rigidly on a robotic servo, which sweeps vertically to cover the third dimension. Fiducial target based extrinsic calibration is applied to acquire transformation matrices between LiDAR and camera. Based on the transformation matrix, we perform registration to fuse the color images from the camera with the 3D points cloud from the LiDAR. The whole system setup has much lower cost as compared to systems using 3D LiDAR and omnidirectional camera. Using pre-calculated transformation matrices instead of feature extraction techniques such as SIFT or SURF in registration gives better fusion result and lower computational complexity. The experiments carried out in office building environment show promising results of our approach.**

***Keywords*—3D indoor modeling, 2D line-scan LiDAR, digital camera, extrinsic calibration, sensor fusion**

I. INTRODUCTION

Building 3D model of environments has become an important task for many applications, such as robot navigation, tourist guidance and cartography. However, it is difficult to construct an accurate 3D model for common consumers, due to the fact of expensive detecting devices, complex system setup and computational costly reconstruction algorithms. In this work, we present a low cost prototype system that enables non-expert users to construct a detailed 3D model of environments in fast, easy approach.

The key problem of building 3D model is how to effectively fuse the data from multiple sensors. The color and texture information are usually collected by camera, and the depth information is captured by range sensor. In order to fuse with 3D points cloud, some existing works take advantage of an omnidirectional camera [1]. We use an entry-level camera for purpose of reducing the cost. The RGB-depth camera like Kinect is sensing device that capture both RGB image and depth image [2, 3]. However, RGB-depth camera can only provide depth information up to a very limited range (around 5m). And its depth estimates are very noisy compare to LiDAR. Therefore, we select LiDAR as our range sensor. There are two different methods for generating 3D points cloud of 3D model from extracted distance information based on LiDAR sensor. Combined 2D line-scan LiDAR with servo [4] to generate 3D points and the density of the 3D points can be changed as the customer needed. The other way is using 3D LiDAR to capture distance information and to generate the 3D points instantly. Since 3D LiDAR is up to 6 times more expensive than 2D line-scan LiDAR, we naturally pick the 2D line-scan LiDAR and mount it on a robotic servo to cover the third dimension.

A prerequisite for fusing data is to extrinsically calibrate the correspondences between camera and LiDAR sensor. A variety of methods exist for extrinsic calibration of imagery and LiDAR coordinate frame. The first sort of methods is proposed to automatically calibrate 2D-3D data without especially multimodal setup. Peter Henry et al. work on a set of corresponding regions to solve a 2D-3D nonlinear shape registration problem based on Levenberg-Marquardt algorithm [5]. In particular, the works of Xiaojin Gong et al. rely on a trihedral configuration [6]. The moving objects are used to find potential transformation between a pair of depth sensors [3]. Harris features [7], SIFT features [2], or SURF features [8] can be used to automatically find corresponding feature pairs from imagery and 3D points cloud. These methods rely on extract special shape or features from two sensors synchronized. It will produce misalignment if the experiment is accomplished under a noisy environment. The second sorts of methods deal with extrinsic calibration to find geometric constraints based on the fiducial targets. A single circle is employed to find the point-to-plane correspondences [9, 10]. Stefano Debattisti et al. rely on a polygonal shape like triangle [11] or diamond [12] to compute extrinsic parameters. These shapes cause range discrepancies for some LiDAR sensors to find corresponding points in camera images. Some algorithms rely on intensity information of LiDAR sensor extracted from a planar checkerboard pattern. Lipu Zhou et al [13] present the convenience of using checkerboard as the calibration target. Nonlinear least squares (NLS) method [6, 11, 14, 15] is employed to calibrate the relative transformation. Random Sample Consensus (RANSAC) algorithm can eliminate the outliers for providing a better estimation to find the correspondences [7, 8, 12, 16, 17]. Raymond Sheh et al. utilize some other methods [4, 18] like linear square method and so on.

In contrast to previous works, this paper proposes a new method to extrinsically calibrate a 3D modeling system with 2D LiDAR and camera. A planar checkerboard pattern is used to extract corners from camera image and intensity image of 2D LiDAR. Instead of finding geometric constraints from 2D-3D data, we rely on the 2D-2D correspondences. A pinhole camera model is applied to project 3D points cloud to 2D plane.

The point-to-point correspondence is estimated based on RANSAC algorithm. To cover 180^o field of view, the camera rotates horizontally in three positions (left, center, right). Thus, we have three transformation matrices between LiDAR and camera in three different poses. After the registration of camera images with 3D points cloud, the transformation matrices are employed in 3D modeling of indoor environments. The experiments carry out in office building corridor area show promising result of our approach.

The outline of the paper is as follows: In Section 2, we give an overview of the LiDAR-camera scan system. Section 3 describes the detail for extrinsic calibration. Section 4 presents the experimental result for indoor 3D modeling. We conclude the paper in Section 5.

II. SYSTEM OVERVIEW

This section describes the design of our proposed system. Figure 1 presents an overview of the 3D modeling procedure. The procedure begins with two types of data: (1) 3D points cloud with intensity information provided by LiDAR (2) color information captured by a digital camera. In order to find the correspondences between 2D-3D data, we project the 3D points cloud to 2D space based on pinhole camera model, and get the 2D intensity image. Next, the checkerboard pattern is extracted from both camera image and LiDAR intensity image. RANSAC algorithm is applied here to refine the geometric constraints detection process. By making use of geometric constraints from the checkerboard pattern, we are able to calculate the extrinsic transformation matrices between camera and LiDAR. After the extrinsic calibration, we use these transformation matrices to match the color images with the intensity images. Afterwards, the 2D points with color information are back projected to 3D for registration of textured 3D points cloud. By aligning 3D color point cloud in different locations, we can finally generate a 3D model of the indoor environment. Moreover, only texting the 3D points can result in some pores inside the 3D modeling, meshing the 3D

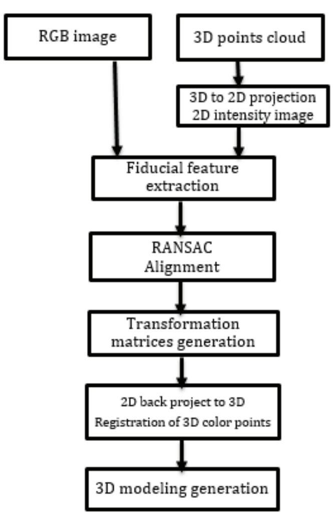


Figure 1: Flow chart for 3D modeling procedure.

realistic color points cloud can provide a more closured 3D modeling.

In this paper, we assemble a LiDAR-camera system for scanning 3D indoor environment. It includes a Hokuyo UTM-30LX-EW 2D line-scan LiDAR, a Turret Widow X servo, and a Nikon S3000 digital camera. As shown in Figure 2, the 2D LiDAR is installed on the servo and the digital camera is rigidly mounted on top of the LiDAR-servo system, the total 3D scanning sensor system is mounted on a pushcart for stop-and-go detection. Figure 3 gives a detailed layout of 3D points cloud generation components – the Hokuyo UTM-30LX-EW line-scan LiDAR, the Turret Widow X servo, the Arduino control board, and the host PC. The Hokuyo LiDAR requires a 12V 1A power supply and communicates with the host PC via the Ethernet interface. The servo moves 2D LiDAR in vertical direction to cover the third dimension. The servo rotates around the pitch axis with 1.05^o elevation angle in 173 steps. Combining with the 2D LiDAR, which has 180^o horizontal field-of-view, the scanning system generates a 3D points cloud covering a half spherical view in front of it. The servo is controlled via a USB connector to the host PC. In order to capture a relatively wide field-of-view, the digital camera rotates in both vertical and horizontal directions. The panoramas are generated through stitching several images captured during vertical rotation of the servo. Figure 4 shows the camera head being moved by the servo in three different horizontal positions, so that the camera can cover the same 180^o horizontal field-of-view as the LiDAR.

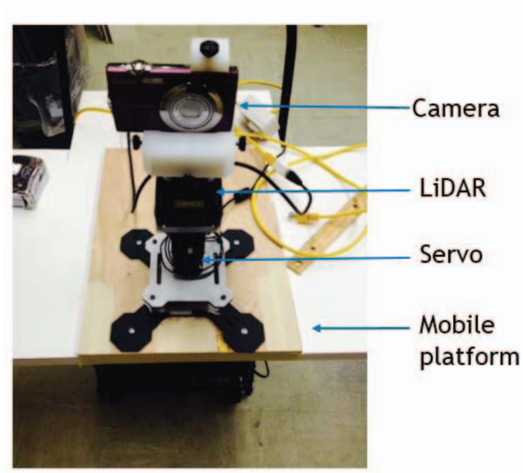


Figure 2: Snapshot of the prototype system setup.

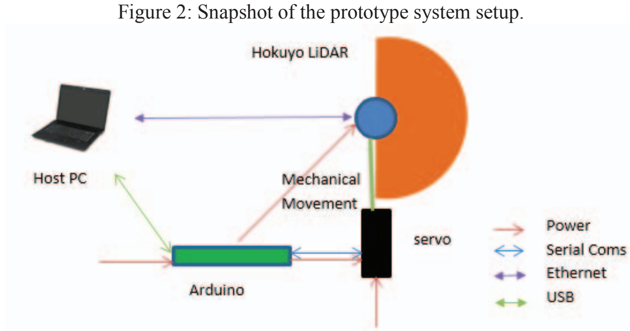


Figure 3: Layout of the LiDAR-servo system.

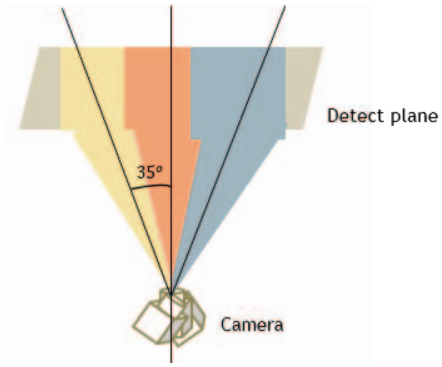


Figure 4: Camera positions for generating three panoramas.

III. EXTRINSIC CALIBRATION

Extrinsic calibration is carried out between a common digital camera connected to the 2D LiDAR-servo system. A fiducial target is used to determine the rigid transformation between camera images and 3D points cloud. A checkerboard containing 7x9 squares with the width of 57mm is employed to find the correspondences between the LiDAR and the camera. Random Sample Consensus (RANSAC) algorithm is applied to eliminate outliers during the calibration process.

As shown in Figure 5, a 3D point in LiDAR calibration plane is represented as $P_l = [x, y, z]^T$ and its related pixel in camera image plane is described as $P_c = [X, Y, 1]^T$. The 3D point P_l with intensity information is projected to a calibration plane under pinhole camera model. The calibration plane is defined at $z = f$ and the projected point in calibration plane is shown as $P = [u, v, 1]^T$. Based on similar triangle rules, we have the relationships as follows:

$$u = f\frac{x}{z}; \qquad v = f\frac{y}{z}; \tag{1}$$

where f is the focal length of the camera. In order to fuse the information from LiDAR and camera, we need to look for the relationship to match P and P_c.

From Figure 6 we can see, after projecting the 3D points to calibration plane, we get a 2D points cloud. These 2D points are interpolated to generate a LiDAR intensity image. Now the problem of extrinsic calibration has become how to find the geometric constraints between a LiDAR intensity image and a camera image using the checkerboard pattern. The transformation of checkerboard pattern from LiDAR calibration coordinate frame to camera coordinate frame is represented by a rigid 3x3 transformation matrix T with 6 degree of freedom.

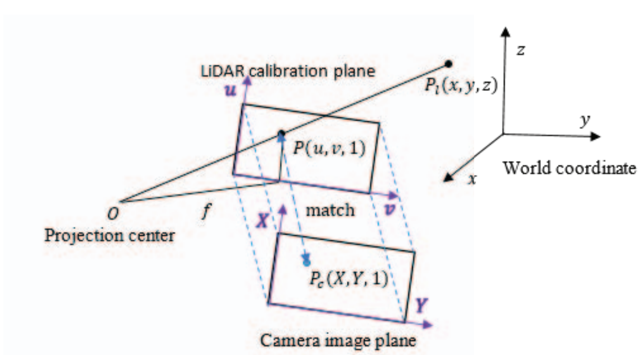


Figure 5: Pinhole camera model.

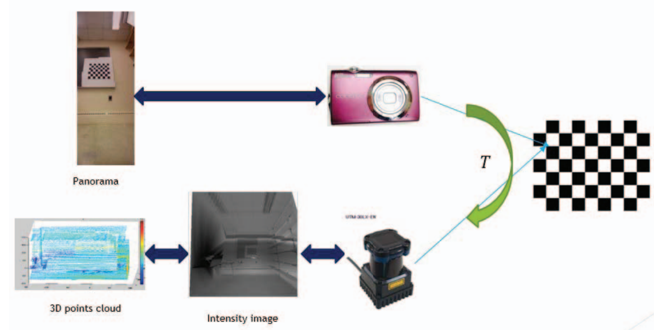


Figure 6: Extrinsic Calibration Procedure.

$$P_c = TP \tag{2}$$

As shown in Figure 7, to obtain the features of checkerboard accurately, we select Region of Interest (ROI) from LiDAR intensity image and camera panorama for checkerboard pattern. Next, we take advantage of Random Sample Consensus (RANSAC) algorithm to eliminate the outliers before estimating the transformation matrix.

The algorithm for generating the transformation matrix is summarized below [16]:

1) Find the inliers for the corners of checkerboard based on RANSAC algorithm.

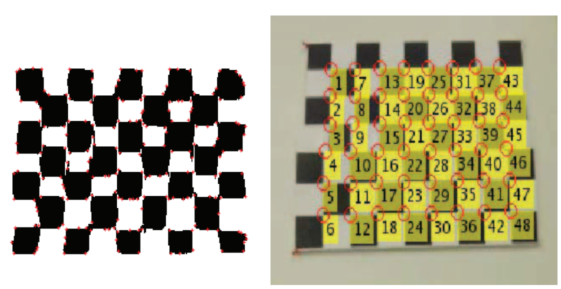
The RANSAC algorithm follows these steps:

a) Randomly select three pairs of points from LiDAR calibration image and camera image for affine transformation.

b) Calculate the transformation matrix T from the selected points.

c) Change the old T, if the distance matrix of the new T is less than the old one.

2) Choose the transformation matrix T which has the maximum inliers.



(a) (b)

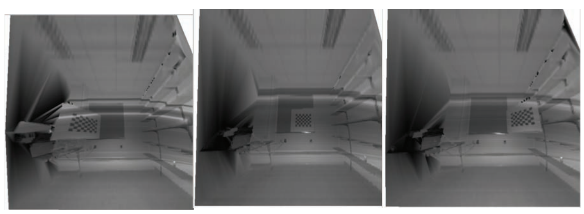
Figure 7: The corners of checkerboard in (a) ROI of LiDAR intensity image (b) ROI of camera panorama.

Use all inlier point pairs to compute a refined transformation matrix T.

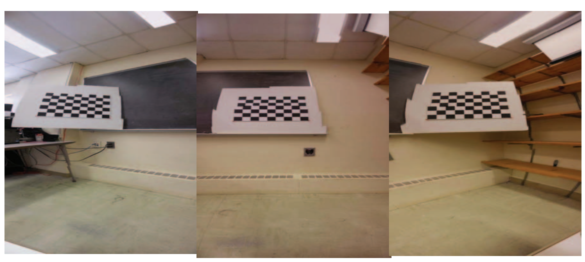
Totally 48 pairs of corner points are selected in calibration process. We get three transformation matrices between the LiDAR and camera in three poses.

We place the checkerboard in three different locations (left, center, and right) to generate camera panoramas and LiDAR intensity images separately. Based on these images, we derive the three transformation matrices. As shown in Figure 8, the color images are seamlessly fused with the intensity images. After generating the transformation matrices, we are able to stitch three camera panoramas together and fuse them with one LiDAR intensity image by applying these transformation matrices. The results are shown in Figure 9. After the fusion, we find the correspondences and assign color value to each LiDAR 2D point.

Finally, we back-project the textured 2D points to 3D color points cloud, as shown in Figure 10. Note that there are dark areas at the border, due to the fact that those points with distance from the projection center shorter than the focal length cannot be projected to the calibration plane.



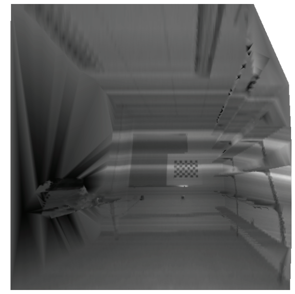
(a)



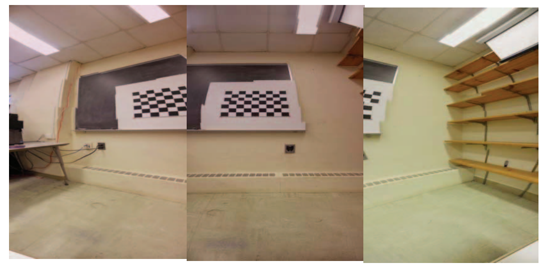
(b)

(c)

Figure 8: Transformation matrices derived process: (a) LiDAR intensity images; (b) Panoramas; (c) Stitched panoramas with intensity images.



(a)



(b)

(c)

Figure 9: Image stitch process: (a) LiDAR intensity image; (b) Panoramas; (c) Stitched three panoramas in one intensity image.

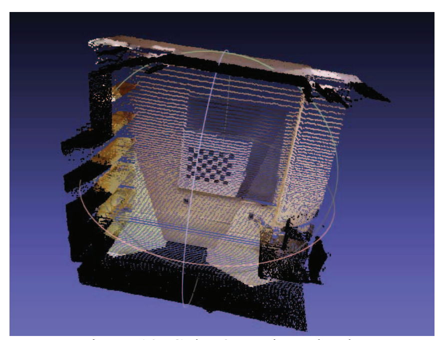


Figure 10: Color 3D points cloud.

IV. EXPERIMENTAL RESULTS

In this section, we apply the extrinsic calibration results to a large indoor environment in an office building corridor. The LiDAR-camera scan system is mounted on a pushcart to record the data in stop-and-go mode. Each scan contains roughly 124,000 LiDAR 3D points. In addition, the camera captures 13 images (640*480 pixels) in every three different poses to form the camera panoramas. In total we

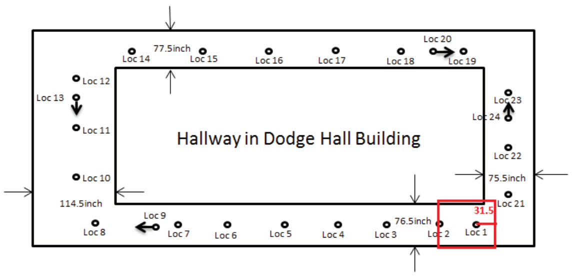


Figure 11: The detecting locations for 3D mapping.

have scanned 24 location. These survey points form a loop in the corridor area as seen in Figure 11.

The measured distance between consecutive scan locations is 150 inches. To colorize the dark areas mentioned in last section, an overlap of 31.5 inches, which is slightly larger than the focal length, is set between consecutive scan locations. For example, in Figure 10, the red rectangle illustrates the scanning region at location 2, which covers the dark areas of location 1. Note that the last scanned location for each hallway is done in the opposite direction (location 9, 13, 20 and 24) to form a closed hallway. By carefully aligning the data of four hallways, we get the final loop shown in Figure 12.

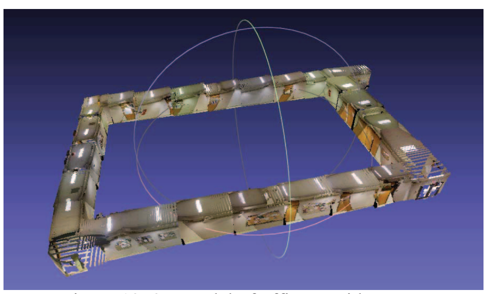


Figure 12: 3D model of office corridor area.

V. CONCLUSION & FUTURE WORK

In this paper, we present a system setup for indoor environment 3D modeling using camera and 2D LiDAR. Our approach utilizes a planar checkerboard pattern in camera-LiDAR extrinsic calibration. Pre-calculated transformation matrices are applied to fuse 3D points cloud with color information captured from the camera. To the best of our knowledge, this system is the first one to achieve low cost, computational efficient and accurate in indoor 3D modeling.

In the future, we will keep working on this prototype system and try to combine motion sensors to the system to enable automatic 3D modeling of indoor environment.

REFERENCES

[1] Gaurav Pandey, James McBride, Silvio Savarese, Ryan Eustice. "Extrinsic calibration of a 3D Laser scanner and an omnidirectional camera," in *7th IFAC Symposium on Intelligent Autonomous Vehicle*, vol. 7 no. 1. 2010.

[2] Hao Du, Peter Henry, Xiaofeng Ren, Marvin Cheng, Dan B Goldman, Steven M. Seitz, Dieter Fox. "Interactive 3D modeling of indoor environments with a consumer depth camera," in *Proceedings of the 13th International Conference on Ubiquitous Computing*, ACM, 2011.

[3] Stephen Miller, Alex Teichman, Sebastian Thrun. "Unsupervised extrinsic calibration of depth sensors in dynamic scenes," In *Processing of IEEE International Conference on Intelligent Robots and Systems*, November 3-7, 2013.

[4] Raymond Sheh, Nawid Jamali, M. Waleed Kadous, Claude Sammut. "A low-cost, compact, lightweight 3D range sensor," in *Proceedings of Australian Conference on Robotics and Automation*, 2006.

[5] Levente Tamas, Robert Frohlich, Zoltan Kato. "Relative pose estimation and fusion of omnidirectional and Lidar camera," in *Proceedings of Computer Vision-ECCV 2014 Workshops,* pp. 640-651, Springer International Publishing, 2012.

[6] Xiaojin Gong, Ying Lin, Jilin Liu. "3D LIDAR-camera extrinsic calibration using an arbitrary trihedron," in *Journal of sensors*, 1 February 2013.

[7] Scott E. Budge, Neeraj S. Badamikar, Xuan Xie. "Automatic registration of fused Lidar/digital imagery (texel images) for three-dimensional image creation," in *Proceedings of Society of Photo-Optical Instrumentation Engineer (SPIE)*, November 2014.

[8] C. Bodensteiner, W. Hubner, K. Jungling, P. Solbrig, M. Arens. "Monocular camera trajectory optimization using LiDAR data," in *Proceedings of the 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, November 2012.

[9] Hatem Alismail, L. Douglas Baker, Brett Browning. "Automatic calibration of a range sensor and camera system," in *Proceedings of the Second Joint 3DIM/3DPVT Conference*, 2012.

[10] Sergio A. Rodriguez F., Vincent Fremont, Philippe Bonnifait. "Extrinsic calibration between a multi-layer lidar and camera," in *Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, August 20-22, 2008.

[11] Stefano Debattisti, Luca Mazzei, Matteo Panciroli. "Automated extrinsic laser and camera inter-calibration using triangular targets," in *Proceedings of IEEE Intelligent Vehicles Symposium, Gold Coast, Australia,* June 23-26 2013.

[12] Yoonsu Park, Seokmin Yun, Chee Sun Won, Kyungeun Cho, Kyhyun Um, Sungdae Sim. "Calibraiton between color camera and 3D LIDAR instruments with a polygonal planar board," in *Journal of sensors*, 2014.

[13] Lipu Zhou, Zhidong Deng. "A new algorithm for extrinsic calibration of a 2D LIDAR and a camera," in *Proceedings of the Measurement Science and Technology*, May 2014.

[14] Qilong Zhang, Robert Pless. "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," in *Proceedings of IEEE International Conference on Intelligent Robots and Systems (IROS)*, vol. 13, 2004.

[15] Faraz M Mirzaei, Dimitrios G Kottas, Stergios I Roumeliotis. "3D LIDAR-camera intrinsic and extrinsic calibration: identifiability and analytical least-squares-based initialization," in the International Journal of *Robotics Research*, vol. 31 no. 4 452-467, April 2012.

[16] Lipu Zhou, Zhidong Deng. "Extrinsic calibration of a camera and a lidar based on decoupling the rotation from the translation," in *Proceedings of Intelligent Vehicle Symposium*, June 3-7, 2012.

[17] Timonthy Liu, Matthew Carlberg, George Chen, Jacky Chen, John Kua, Avideh Zakhor. "Indoor localization and visualization using a human-operated backpack system," in *Proceedings of International Conference on Indoor Position and Indoor Navigation (IPIN),* September 15-17, 2010.

[18] Gaurav Pandey, James R. MaBride, Silvio Savarese, Ryan M. Eustice. "Automatic targetless extrinsic calibration of a 3D Lidar and camera by maximizing mutual information," in *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, July 2012.