

# SLAT-Calib: Extrinsic Calibration between a Sparse 3D LiDAR and a Limited-FOV Low-resolution Thermal Camera

Jun Zhang<sup>1\*</sup>, Ran Zhang<sup>1\*</sup>, Yufeng Yue<sup>1</sup>, Chule Yang<sup>1</sup>, Mingxing Wen<sup>1</sup> and Danwei Wang<sup>1</sup>

**Abstract**—Accurate estimation of extrinsic parameter (rotation matrix and translation vector) between heterogeneous sensors is important for fusing complementary information. However, the extrinsic calibration between a sparse 3D LiDAR and a thermal camera is challenging, mainly because of the difficulties to accurately extract common features from a sparse point cloud and a thermal image which has limited-FOV and low-resolution. Previous methods either rely on a dense depth sensor or a visual camera to facilitate the feature extraction process. To address the problem, SLAT-Calib (Sparse Lidar And Thermal camera Calibration) is proposed. By observing that circular holes could be detected from both sensors, a specially designed calibration board (a rectangular board with four circular holes) is introduced. Four circle centers in 3D space are used as common features. The benefit is point features are accurate and reliable for feature matching. To extract four circle centers from the thermal camera, three steps are carried out: First, a method is proposed to accurately detect the four circles. Then, the homography matrix of the calibration board can be figured out. Lastly, 3D coordinates of the circle centers are calculated by decomposing the homography matrix. From the LiDAR frame, the four circle centers can be segmented out as long as two laser beams pass through each circle. At last, optimal extrinsic parameter is calculated by minimizing the matching error between the four pairs of 3D circle centers. Quantitative and qualitative experiments are carried out. In simulation, SLAT-Calib outperforms two methods by a large margin. In real environment, it achieves a re-projection error (RMSE) of 0.62 pixel.

## I. INTRODUCTION

In challenging environments, the fusion of a sparse 3D LiDAR and a thermal camera could help robots to enhance the perception ability. For example, in low-illumination environments, visual cameras will fail to provide texture information, but a thermal camera still works properly. In addition, to obtain accurate depth information, commonly used RGB-D cameras can only work within a limited range (5m), but a 3D LiDAR has a long detection range (up to hundred meters). To achieve the fusion, accurate extrinsic calibration is indispensable, with which the transformation matrix between the sensor frames could be obtained. Thus, the depth and temperature information provided by each sensor could be integrated. The integrated information provides obvious advantages and benefits for various scenarios, such

\* This research was partially supported by the ST Engineering - NTU Corporate Lab through the NRF corporate lab@university scheme. (*Corresponding author: Yufeng Yue*)

<sup>1</sup>Jun Zhang, Ran Zhang, Yufeng Yue, Chule Yang, Mingxing Wen and Danwei Wang are with School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore {jzhang061, rzhang017, yyue001, yang0438, mingxing001}@e.ntu.edu.sg, edwwang@ntu.edu.sg

\* Co-first authorship.

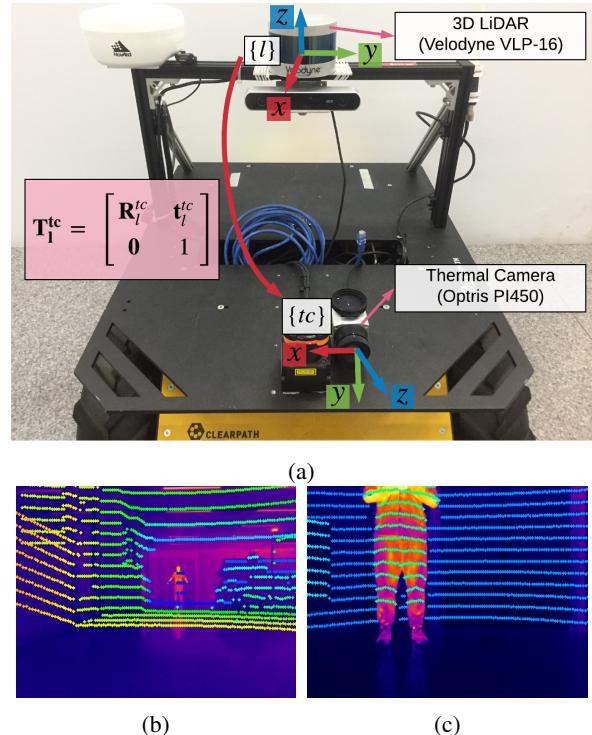


Fig. 1: To obtain the extrinsic parameter (transformation matrix  $T_l^{tc}$ ) between a Sparse 3D LiDAR And a Thermal camera, an off-line one-step method SLAT-Calib is proposed. (a) The sensor frames of LiDAR and thermal camera are represented as  $\{l\}$  and  $\{tc\}$ ,  $x$ -,  $y$ -,  $z$ - axis are shown in red, green and blue color. The transformation matrix  $T_l^{tc}$  consists of a rotation matrix  $R_l^{tc}$  and a translation vector  $t_l^{tc}$ . (b)(c) Project 3D point cloud from LiDAR to undistorted thermal image (human at about 14m and 2.5m).

as thermal mapping [1], [2], thermal reconstruction [3], and human 3D position estimation at night [4].

However, the extrinsic calibration between a sparse 3D LiDAR and a thermal camera is challenging. The difficulty mainly comes from the heterogeneous characteristics of the two sensors, which makes it hard to extract common features. In general, to work properly, thermal camera requires the existence of temperature difference. Thus, the commonly used chessboard [5], [6], [7] cannot be used for thermal camera. In addition, sparse 3D LiDAR generates a sparse point cloud, to extract features from sparse point cloud is far more challenging than from a dense point cloud. To solve the problems, previous methods either rely on a very dense 3D LiDAR [2], [8] to extract features from point cloud, or introduce a visual camera as the medium [9] to assist feature extraction.

In this paper, a novel method SLAT-Calib is proposed. The main novelty is, by observing that the centers of circular holes can be detected from both sensors, a specially designed calibration board (rectangular board with four circular holes, Fig. 3) is introduced for calibration. Furthermore, an algorithm is proposed to accurately and robustly calculate the 3D coordinates of the four circle centers from a thermal image. That is very important to enable the method to work with a monocular camera. Meanwhile, the four 3D circle centers could be segmented out from a sparse 3D LiDAR, as long as two laser beams pass through each circle.

SLAT-Calib is a one-step method and differs greatly from our previous two-step method [9]. SLAT-Calib is inspired by [10], however, [10] relies on a stereo camera. In contrast, SLAT-Calib is designed for monocular camera.

This paper makes the following contributions:

- As far as we know, SLAT-Calib is the first one-step method to obtain the extrinsic parameter between a sparse 3D LiDAR and a thermal camera.
- This paper provides new directions to accurately and reliably extract common features from a sparse 3D LiDAR and a monocular camera.
- SLAT-Calib achieves a re-projection error (RMSE) of 0.62 pixel in real environment.

The paper is organized as follows: Related works are reviewed in Section II. The proposed method SLAT-Calib is explained in Section III. Qualitative and quantitative experiments are performed in Section IV. At last, Section V concludes the paper and discusses future work.

## II. LITERATURE REVIEW

As mentioned above, to calibrate a sparse 3D LiDAR and a thermal camera, the main challenge is twofold: sparseness of the 3D LiDAR and thermal camera. Therefore, the literature review will be done along these two directions.

### 1) Extrinsic Calibration Involves a Sparse 3D LiDAR:

First, the extrinsic calibration between a LiDAR and a camera can be traced back to [11]. It introduces chessboard into the calibration between a 2D LiDAR and a camera. The work is extended to a 3D LiDAR and a camera by [12]. To simplify the calibration process, [5] proposes to use multiple checkerboards and achieves the intrinsic and extrinsic calibration between a 3D LiDAR and multiple cameras in one shot. They also made the calibration process convenient by providing an on-line calibration toolbox [13]. To alleviate the cumbersome of preparing multiple checkerboards, a new calibration board - a rectangular board with four circular holes - is proposed in [14]. However, those methods are designed for a dense 3D LiDAR (e.g. Velodyne HDL-32E or HDL-64E). For sparse 3D LiDAR (e.g. Velodyne VLP-16), the performance is unsatisfactory.

Some research is carried out to solve the problem of LiDAR sparseness. Based on the newly proposed calibration board in [14], [10] makes a significant progress by proposing a method to extract the four 3D circle centers from a sparse 3D LiDAR. However, [10] requires a stereo camera. The constraint of a stereo camera is released by [6]. They

observed that by tilting a chessboard, the chessboard can be detected from both sensors, even if it is a sparse LiDAR. Following that, [7] improves the accuracy by using 3D point and plane as features. Even though these methods work well with sparse 3D LiDAR, they were designed for visual cameras, and could not work with a thermal camera.

2) *Extrinsic Calibration Involves a Thermal Camera:* To calibrate a 3D LiDAR with a thermal camera, the existing literatures rely on a dense depth sensor. The calibration between a 3D LiDAR and a thermal camera is achieved by [8], however, the method adopts a rotating 2D LiDAR to produce very dense 3D point cloud. Following that, the method is improved by utilizing a new calibration board - a rectangular board with one large circular hole [2]. However, to detect the circle, the algorithm still relies on a dense point cloud.

In addition, the extrinsic calibration between a structured-light depth sensor and a thermal camera can be found in [1]. It also requires a dense point cloud to extract non-collinear straight lines from both sensors. Our previous work achieves the calibration between a sparse 3D LiDAR and a thermal camera in two steps [9]. However, a visual camera is introduced as the medium, thus the calibration error increases.

Meanwhile, to recover 3D coordinates of a chessboard from 2D image, Zhang's method [15] provides the theory basis. Even it is proposed for a visual camera, it is an important reference for thermal camera.

In conclusion, none of the methods could work well with the configuration of a sparse 3D LiDAR (Velodyne VLP-16) and a thermal camera. Therefore, inspired by [10] and [15], an one-step method SLAT-Calib is proposed to accurately obtain the extrinsic parameter.

## III. METHODOLOGY

In this section, the details of SLAT-Calib is explained. SLAT-Calib is inspired by [10], however, [10] relies on a stereo camera, SLAT-Calib is designed for monocular camera. The main difference resides in the approach to extract four 3D circle centers from the camera: while [10] extracts the circle centers from a 3D point cloud generated by a stereo camera, SLAT-Calib takes advantage of the homography matrix to calculate 3D circle centers from one image. The flowchart of SLAT-Calib is shown in Fig. 2.

### A. Notations

1) *Extrinsic Parameter:* As shown in Fig. 1a, the two sensor frames are represented as  $\{l\}$  (LiDAR frame) and  $\{tc\}$  (thermal camera frame). Generally, the objective of extrinsic calibration is to estimate the unknown transformation matrix  $T_l^{tc}$  between the two sensor frames. With  $T_l^{tc}$ , a 3D point  $P_l = (X_l, Y_l, Z_l)^T$  in the LiDAR frame can be transformed to a 3D point  $\hat{P}_l = (\hat{X}_l, \hat{Y}_l, \hat{Z}_l)^T$  in the thermal camera frame by Eq. (1).

$$\begin{bmatrix} \hat{P}_l \\ 1 \end{bmatrix} = T_l^{tc} \begin{bmatrix} P_l \\ 1 \end{bmatrix} \quad (1)$$

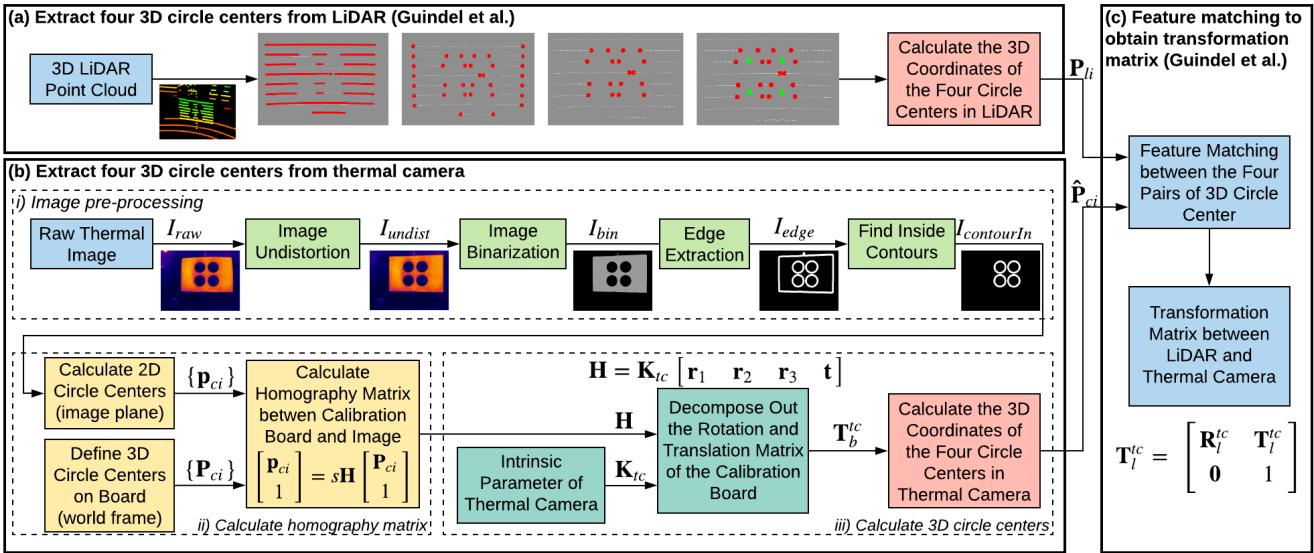


Fig. 2: Flowchart of the proposed method SLAT-Calib. (a) To extract four 3D circle centers from a sparse 3D LiDAR, the method proposed in [10] is adopted. (b) To extract four 3D circle centers from a monocular thermal camera, three steps are performed: i) Image pre-processing - to detect the four circles from thermal image. ii) Calculate homography matrix - to figure out the homography matrix  $\mathbf{H}$  between the calibration board and the thermal image. iii) Calculate 3D circle centers - the 3D coordinates of the four circle centers can be calculated by utilizing the intrinsic parameter  $\mathbf{K}_{tc}$  of the thermal camera and the homography matrix  $\mathbf{H}$ . (c) To match the 3D circle centers and output the final transformation matrix, the method proposed in [10] is utilized.

$\mathbf{T}_l^{tc} \in SE(3)$  consists of two parts: a rotation matrix  $\mathbf{R}_l^{tc} \in SO(3)$  and a translation vector  $\mathbf{t}_l^{tc} \in \mathbb{R}^3$  (Eq. (2)).

$$\mathbf{T}_l^{tc} = \begin{bmatrix} \mathbf{R}_l^{tc} & \mathbf{t}_l^{tc} \\ \mathbf{0} & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4} | \mathbf{R}_l^{tc} \in SO(3), \mathbf{t}_l^{tc} \in \mathbb{R}^3 \quad (2)$$

2) *Intrinsic Parameter*: In this paper, it is assumed that the intrinsic and distortion parameter  $\mathbf{K}_{tc}$  and  $\mathbf{D}_{tc}$  of the thermal camera are known (our previous two-step method [9] can be referred). Using  $\mathbf{K}_{tc}$  and  $\mathbf{D}_{tc}$ , the transformed point  $\hat{\mathbf{P}}_l$  can be projected to  $\hat{\mathbf{p}}_l = (\hat{u}_l, \hat{v}_l)^T$  on undistorted thermal image with Eq. (3).

$$\begin{bmatrix} \hat{\mathbf{p}}_l \\ 1 \end{bmatrix} = \frac{1}{\hat{Z}_l} \mathbf{K}_{tc} \hat{\mathbf{P}}_l \quad (3)$$

### B. Extract Four 3D Circle Centers From Thermal Camera

This part is the main novelty to make [10] work with a monocular camera, rather than a stereo camera. It mainly consists of four steps: selection of the calibration board, image pre-processing, calculate homography matrix, and calculate 3D circle centers from a thermal image.

1) *Selection of Calibration Board*: The black and white chessboard is commonly used to calibrate visual cameras [15]. However, it cannot work well with a thermal camera, because the corner points could not be "seen" properly by the thermal camera. Fortunately, we found that a specially designed rectangular board [10] with four circular holes (Fig. 3a) could be "seen" properly by the thermal camera. The four circles and circle centers could be accurately detected on the thermal image (Fig. 3b). Meanwhile, this board is proved to work well with a sparse 3D LiDAR (e.g., Velodyne VLP-16) in [10]. Therefore, this rectangular board is selected in this paper. The board is made from acrylic material by a laser cutting machine. To heat up the board, [9] can be referred.

2) *Image Pre-processing*: Several pre-processing steps on the thermal image are carried out, as shown in Fig. 2b(i).

First, the raw thermal image  $I_{raw}$  is undistorted to  $I_{undist}$ , using the intrinsic and distortion parameter  $\mathbf{K}_{tc}$  and  $\mathbf{D}_{tc}$ . Then, the undistorted image is binarized to  $I_{bin}$ . The threshold to obtain  $I_{bin}$  is  $\sigma_{bin} = (60, 155)$ . After passing through a canny edge detector, the edge image  $I_{edge}$  is acquired. At last, to obtain the four circles, the contour detection algorithm [16] is utilized. The pseudo-code of image pre-processing step is shown in Algorithm 1 line 1 - 8.

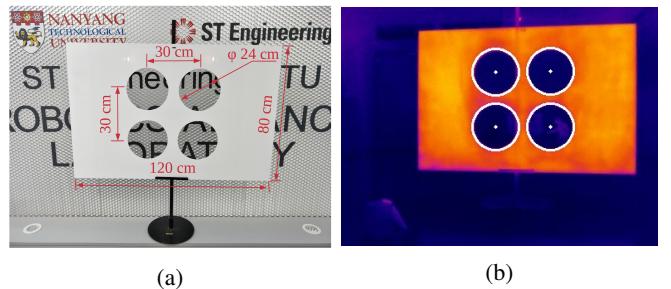


Fig. 3: (a) The selected calibration board and the size. (b) The detected circle contours and circle centers are drawn on the undistorted thermal image (shown in white color).

3) *Calculate Homography Matrix*: Homography matrix  $\mathbf{H}$  builds the mapping relationship between two correspondent point sets on two planes. In this paper, the two correspondent point sets are: the four circle centers detected on the image plane  $\mathbb{I}_c = \{\mathbf{p}_{ci} = (u_{ci}, v_{ci})^T | i = 1, 2, 3, 4\}$  (unit: pixel); the four circle centers defined on the calibration board plane  $\mathbb{P}_c = \{\mathbf{P}_{ci} = (X_{ci}, Y_{ci}, Z_{ci})^T | i = 1, 2, 3, 4\}$  (unit: m). As shown in Fig. 2b(ii), the pixel coordinate  $\mathbf{p}_{ci}$  of the four circle centers can be calculated with the image  $I_{contourIn}$ . At the same time, because the size of the calibration board

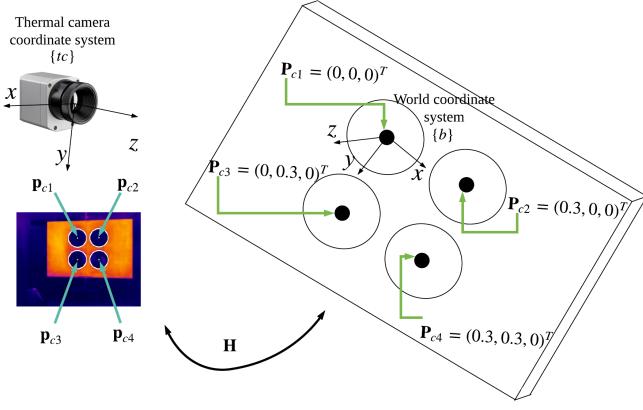


Fig. 4: To calculate homography matrix  $\mathbf{H}$ , pixel coordinate of the four circle centers  $\mathbf{p}_{ci}$  (unit: pixel) is calculated and the coordinate of the four circle centers in the world coordinate system  $\{b\}$  is defined as  $\mathbf{P}_{ci}$  (unit: m), using the known size of the board.

is known,  $\mathbf{P}_{ci}$  can be defined as shown in Fig. 4. World coordinate system is defined on the board plane, represented as  $\{b\}$ , the origin point is the top left circle center, the  $z$ -axis is perpendicular to the plane. It is obvious that for all  $\mathbf{P}_{ci}$ , the  $z$ -coordinate  $Z_{ci} = 0$ .

The relationship between  $\mathbf{p}_{ci}$  and  $\mathbf{P}_{ci}$  can be established with homography matrix by Eq. (4).  $[\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{t}]$  is the rotation and translation between the calibration board  $\{b\}$  and the thermal camera  $\{tc\}$ .  $s$  is a scaling factor, thus  $\mathbf{H}$  has 8 degree of freedom. One pair of point  $(\mathbf{p}_{ci}, \mathbf{P}_{ci})$  can provide two equations, thus at least four pairs of point can solve out the homography matrix  $\mathbf{H}$ . In this paper, it is calculated via Singular Value Decomposition (SVD) [15]. The pseudo-code of calculating homography matrix is shown in Algorithm 1 line 9 - 16.

$$\begin{bmatrix} u_{ci} \\ v_{ci} \\ 1 \end{bmatrix} = s \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X_{ci} \\ Y_{ci} \\ 0 \\ 1 \end{bmatrix}$$

$$= s \mathbf{K}_{tc} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X_{ci} \\ Y_{ci} \\ 1 \end{bmatrix}$$

$$= s \mathbf{H} \begin{bmatrix} X_{ci} \\ Y_{ci} \\ 1 \end{bmatrix} \quad (4)$$

#### 4) Calculate 3D Circle Centers in Thermal Camera:

As shown in Fig. 2b(iii), once the homography matrix  $\mathbf{H}$  is obtained and the intrinsic parameter  $\mathbf{K}_{tc}$  is known,  $[\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{t}]$  can be decomposed out from  $\mathbf{H}$ , using the closed-form solution proposed in Zhang's method [15]. Thus, the transformation matrix  $\mathbf{T}_b^{tc}$  can be calculated as Eq. (5). Then, the four circle centers  $\mathbf{P}_{ci}$  in the world frame  $\{b\}$  can be transformed to the thermal camera frame  $\{tc\}$  by Eq. (6). The transformed points  $\hat{\mathbf{P}}_{ci}$  ( $i = 1, 2, 3, 4$ ) are the extracted four 3D circle centers from the thermal camera. The pseudo-code for calculating 3D circle centers in thermal camera is

**Algorithm 1** Extract four 3D circle centers from monocular thermal camera

---

```

1: procedure PREPROCESSING( $I_{raw}$ )
2:   // Undistortion, binarization, and edge extraction
3:    $I_{undist} \leftarrow undistort(I_{raw})$ 
4:    $I_{bin} \leftarrow binarization(I_{undist})$ 
5:    $I_{edge} \leftarrow cannyEdge(I_{bin})$ 
6:   // Find and filter contours
7:    $I_{contourIn} \leftarrow findAndFilterContours(I_{edge})$ 
8: end procedure

9: procedure CALCULATE HOMOGRAPHY
10:  // Define the 3D coordinates of the four circle centers
11:   $\mathbb{P}_c = \{\mathbf{P}_{ci}\}$  in the world frame  $\{b\}$ 
12:   $\mathbb{P}_c \leftarrow setWorldCoord()$ 
13:  // Calculate the 2D pixel coordinates of the four
14:  // circle centers  $\mathbb{I}_c = \{\mathbf{p}_{ci}\}$  in the image  $I_{contourIn}$ 
15:   $\mathbb{I}_c \leftarrow calculateImageCoord(I_{contourIn})$ 
16:  // Calculate homography matrix  $\mathbf{H}$ 
17:   $\mathbf{H} \leftarrow calculateHomography(\mathbb{P}_c, \mathbb{I}_c)$ 
18: end procedure

17: procedure CALCULATE 3D CENTERS( $\mathbf{H}, \mathbb{P}_c, \mathbf{K}_{tc}$ )
18:  // Calculate transformation matrix  $\mathbf{T}_b^{tc}$  from the cal-
19:  // ibration board to the thermal camera
20:   $\mathbf{T}_b^{tc} \leftarrow decomposeHomography(\mathbf{K}_{tc})$ 
21:  // Transform the four 3D circle centers  $\mathbf{P}_{ci}$  into the
22:  // thermal camera frame  $\{tc\}$ 
21:   $\hat{\mathbf{P}}_{ci} \leftarrow transformPoints(\mathbf{P}_{ci}, \mathbf{T}_b^{tc})$ 
22: end procedure

```

---

shown in Algorithm 1 line 17 - 22.

$$\mathbf{T}_b^{tc} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$\begin{bmatrix} \hat{\mathbf{P}}_{ci} \\ 1 \end{bmatrix} = \mathbf{T}_b^{tc} \begin{bmatrix} \mathbf{P}_{ci} \\ 1 \end{bmatrix} \quad (6)$$

#### C. Extract Four 3D Circle Centers from Sparse 3D LiDAR

To extract four 3D circle centers  $\mathbf{P}_{li}$  ( $i = 1, 2, 3, 4$ ) from the sparse 3D LiDAR, the same algorithm presented in [10] is adopted. Since the detection method in [10] requires at least two laser beams pass through each circles, a suitable distance to put the board is important. Through testing, it is found that the proper distance range is  $[1.5m, 3.2m]$ . The raw point cloud of the calibration board is shown in Fig. 5a. The extracted four 3D circle centers from LiDAR  $\mathbf{P}_{li}$  are shown as green points in Fig. 5b.

At last, the optimal transformation matrix  $\mathbf{T}_l^{tc}$  can be calculated by minimizing the matching error between  $\hat{\mathbf{P}}_{ci}$  and  $\mathbf{P}_{li}$  ( $i = 1, 2, 3, 4$ ) [10], as shown in Fig. 2c.

## IV. EXPERIMENTAL RESULTS

### A. Experimental Platform

The experimental platform is shown in Fig. 1a. The thermal camera is Optris PI450 (resolution:  $382 \times 288$ , FOV:  $53^\circ(H) \times 42^\circ(V)$ ). The sparse 3D LiDAR is Velodyne

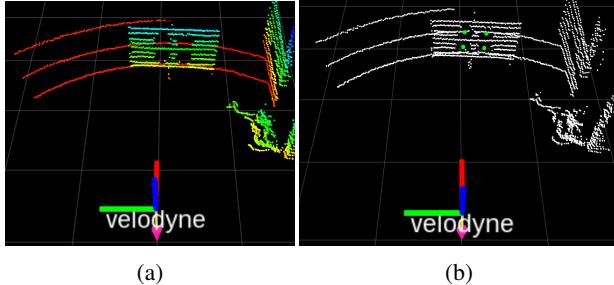


Fig. 5: (a) The raw point cloud collected with the sparse 3D LiDAR Velodyne VLP-16. The calibration board is at about 2.5m relative to the LiDAR frame. (b) The extracted four 3D circle centers from LiDAR  $\mathbf{P}_{li}$  (shown as green points).

VLP-16 (FOV:  $360^\circ(H) \times 30^\circ(V)$ , resolution:  $0.1^\circ\text{-}0.4^\circ(H)$ ,  $2^\circ(V)$ ). The sensors are mounted on a robot (Clearpath HUSKY A200). All algorithms are running on a laptop (Dell Precision M2800), with Intel® Core™ i7-4610M CPU @ 3.00GHz  $\times$  4, 16GB RAM, Ubuntu 16.04, and ROS kinetic.

### B. Quantitative Analysis

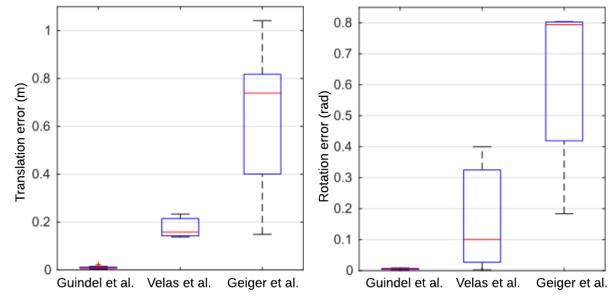
1) *Compare with Ground Truth in Simulation Environment:* The simulation environment is provided by [10] as Gazebo. In simulation environment, the ground truth of the extrinsic parameter  $\mathbf{R}_g, \mathbf{t}_g$  can be obtained. Therefore, the calibration accuracy can be evaluated by comparing the calibration result  $\mathbf{R}, \mathbf{t}$  with the ground truth. To calculate calibration error, the method presented in [5] is adopted: the translation error  $e_t$  is calculated via Eq. (7) and the rotation error  $e_r$  is calculated via Eq. (8).

SLAT-Calib (Ours) is compared with [10] (Guindel et al.), [14] (Velas et al.) and [5] (Geiger et al.). However, these methods are designed for visual camera, rather than thermal camera. Specifically, Guindel et al. requires a stereo visual camera, Velas et al. and Geiger et al. can work with a monocular visual camera. To compare fairly in Gazebo, the same camera presented in [10] is adopted, thermal image is simulated by gray-scaling the rgb image.

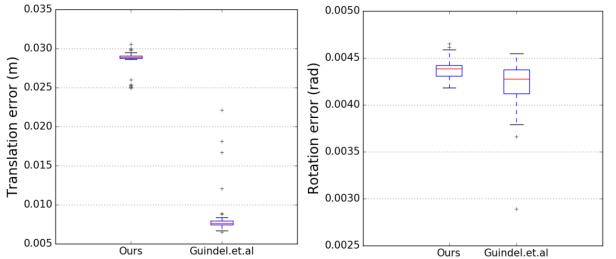
Using the same configuration of simulation environment as [10], the translation and rotation error are compared among Ours, Guindel et al., Velas et al. and Geiger et al. The result is shown in Fig. 6. It can be observed that: From Fig. 6a, Guindel et al. shows much better accuracy than the two monocular camera-based methods; From Fig. 6b, our method achieves comparable accuracy with Guindel et al.; More importantly, combine both Fig. 6a and Fig. 6b, the accuracy can be compared among the three monocular camera-based methods - Ours, Velas et al. and Geiger et al. For translation error  $e_t$ , Ours is no more than 0.03m, Velas et al. is larger than 0.1m, and Geiger et al. is larger than 0.7m. For rotation error  $e_r$ , Ours also shows the best accuracy.

$$e_t = \|\mathbf{t} - \mathbf{t}_g\| \quad (7)$$

$$e_r = \angle(\mathbf{R}^{-1}\mathbf{R}_g) \quad (8)$$



(a) The result presented in [10].



(b) Ours vs Guindel et al.

Fig. 6: Compare translation error  $e_t$  and rotation error  $e_r$  in simulation (using the same settings as [10]). Our method (monocular camera based) achieves comparable accuracy with Guindel et al. (stereo camera based); More importantly, our method shows the best accuracy among the three monocular camera-based methods.

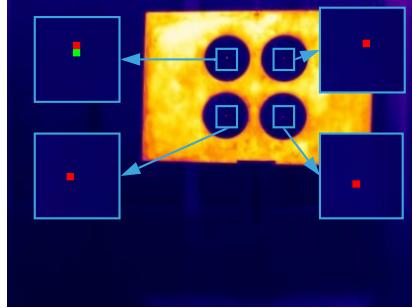
2) *Compute Re-projection Error (RMSE) in Real Environment:* In real environment, the ground truth cannot be obtained, thus re-projection error is utilized to evaluate the calibration accuracy. To calculate the re-projection error, first, the four 3D circle centers detected in LiDAR  $\{\mathbf{P}_{li}\}$  are projected to  $\{\hat{\mathbf{p}}_{li}\}$  on the undistorted thermal image, using Eq. (1) and (3). Then, the re-projection error Root Mean Square Error (RMSE) is calculated using Eq. (9) ( $N = 4$ ).

Since the distance of the calibration board to the thermal camera affects the calibration accuracy, the calibration board is put at different distances to figure out a proper distance. The board is put at 1.7m, 2.0m, 2.5m, 3.0m. The corresponding re-projection error is listed in Table I. It can be observed that to put the board at 2.5m achieves the best performance. The average re-projection error (RMSE) is the smallest: 0.62 pixel. The covariance is also the smallest: 0.0077 pixel.

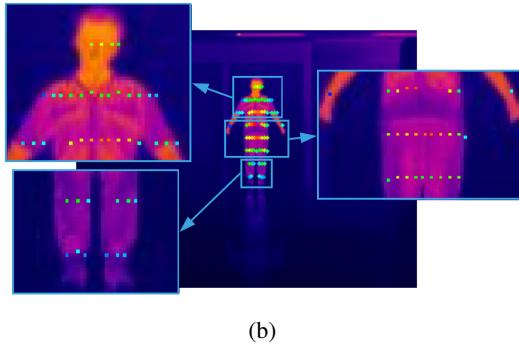
$$e_{reproj} = \sqrt{\frac{\sum_{i=1}^N (\|\hat{\mathbf{p}}_{li} - \mathbf{p}_{ci}\|_2)^2}{N}} \quad (9)$$

### C. Qualitative Analysis

Qualitative analysis can be done by visualizing the re-projection error  $e_{reproj}$  (Fig. 7a). Green pixels (squares) are  $\mathbf{p}_{ci}$ , the circle centers detected from thermal image. Red pixels (squares) are  $\hat{\mathbf{p}}_{li}$ , the circle centers detected from 3D LiDAR and projected onto thermal image. It can be observed that there is only 1 pixel error for the top left center, and 0 pixel error for the other three centers.



(a)



(b)

Fig. 7: Qualitative analysis of SLAT-Calib. (a) Zoom in to visualize the re-projection error  $e_{reproj}$  (board at 2.5m). (b) Zoom in to visualize the projection error of a human at 6m.

The accuracy can be verified by observing an object (such as a human) at different distance. As shown in Fig. 1b, 1c and 7b, the human is at about 14m, 2.5m and 6m, respectively. Those discrete colorful points are the 3D LiDAR points projected onto undistorted thermal image. It can be observed that, the projected points correctly falls into the corresponding pixel points. That is to say, even the calibration is done at 2.5m, SLAT-Calib works well for further distances (14m and 6m).

## V. CONCLUSIONS AND FUTURE WORK

Accurate extrinsic calibration between a sparse 3D LiDAR and a monocular, limited-FOV, low-resolution thermal camera is challenging. It is because of the difficulties to accurately and reliably extract common features. Thus, a novel method SLAT-Calib is proposed in this paper. As far as we know, this is the first one-step method to solve this problem. First, by observing that circular holes can be detected by both sensors, a specially designed calibration board (a rectangular board with four circular holes) is introduced. Second, a method is proposed to accurately calculate the 3D coordinates of the four circle centers from a monocular thermal camera. Quantitative and qualitative experiments demonstrated the accuracy and generality of SLAT-Calib. In simulation, SLAT-Calib achieves comparable accuracy with a stereo camera-based method. More importantly, SLAT-Calib outperforms two monocular camera-based methods by a large margin. In real environment, SLAT-Calib achieves a re-projection error (RMSE) of 0.62 pixel. Meanwhile, it is

TABLE I. RE-PROJECTION ERROR RMSE  $e_{reproj}$  OF SLAT-CALIB. THE CALIBRATION BOARD IS PUT AT DIFFERENT DISTANCE (IN REAL ENVIRONMENT).

Calibration Board Distance (m)	Max (pixel)	Min (pixel)	Mean (pixel)	Cov. (pixel)
1.7	1.56	1.16	1.29	0.0099
2.0	1.36	<b>0.41</b>	0.77	0.0282
2.5	<b>0.82</b>	0.48	<b>0.62</b>	<b>0.0077</b>
3.0	2.03	1.51	1.84	0.0420

shown that SLAT-Calib can be used not only with thermal cameras, but also with visual cameras. In the future, the influence of the orientation of the calibration board on the accuracy is worth to be explored.

## REFERENCES

- [1] S. Vidas, P. Moghadam, and M. Bosse, “3d thermal mapping of building interiors using an rgb-d and thermal camera,” in *2013 IEEE International Conference on Robotics and Automation*, pp. 2311–2318, May 2013.
- [2] A. K. Krishnan and S. Saripalli, “Cross-calibration of rgb and thermal cameras with a lidar for rgb-depth-thermal mapping,” *Unmanned Systems*, vol. 05, no. 02, pp. 59–78, 2017.
- [3] Y. Cao, B. Xu, Z. Ye, J. Yang, Y. Cao, C.-L. Tisse, and X. Li, “Depth and thermal sensor fusion to enhance 3D thermographic reconstruction,” *Opt. Express*, vol. 26, pp. 8179–8193, apr 2018.
- [4] Y. Yue, C. Yang, J. Zhang, M. Wen, B. Dai, and D. Wang, “Simultaneous collaborative mapping and reasoning in dynamic unstructured environments,” in *ICRA 2019 Workshop on Robot Teammates Operating in Dynamic, Unstructured Environments (RT-DUNE)*, 2019.
- [5] A. Geiger, F. Moosmann, Ö. Car, and B. Schuster, “Automatic camera and range sensor calibration using a single shot,” in *2012 IEEE International Conference on Robotics and Automation*, pp. 3936–3943, May 2012.
- [6] L. Zhou, Z. Li, and M. Kaess, “Automatic extrinsic calibration of a camera and a 3d lidar using line and plane correspondences,” in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5562–5569, Oct 2018.
- [7] S. Verma, J. S. Berrio, S. Worrall, and E. M. Nebot, “Automatic extrinsic calibration between a camera and a 3d lidar using 3d point and plane correspondences,” *ArXiv*, vol. abs/1904.12433, 2019.
- [8] A. K. Krishnan, B. Stinnett, and S. Saripalli, “Cross-calibration of rgb and thermal cameras with a lidar,” in *IROS 2015 Workshop on Alternative Sensing for Robot Perception*, 2015.
- [9] J. Zhang, P. Siritanawan, Y. Yue, C. Yang, M. Wen, and D. Wang, “A two-step method for extrinsic calibration between a sparse 3d lidar and a thermal camera,” in *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pp. 1039–1044, Nov 2018.
- [10] C. Guindel, J. Beltrán, D. Martín, and F. García, “Automatic extrinsic calibration for lidar-stereo vehicle sensor setups,” in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–6, Oct 2017.
- [11] Q. Zhang and R. Pless, “Extrinsic calibration of a camera and laser range finder (improves camera calibration),” in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, vol. 3, pp. 2301–2306 vol.3, Sept 2004.
- [12] R. Unnikrishnan and M. Hebert, “Fast extrinsic calibration of a laser rangefinder to a camera,” Tech. Rep. CMU-RI-TR-05-09, Carnegie Mellon University, Pittsburgh, PA, July 2005.
- [13] A. Geiger, F. Moosmann, Ö. Car, and B. Schuster, “Camera and range sensor calibration toolbox,” 2012. Available at <http://www.cvlibs.net/software/calibration>.
- [14] M. Velas, M. Spanel, Z. Materna, and A. Herout, “Calibration of rgb camera with velodyne lidar,” in *WSCG 2014*, 2014.
- [15] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1330–1334, Nov 2000.
- [16] S. Suzuki and K. be, “Topological structural analysis of digitized binary images by border following,” *Computer Vision, Graphics, and Image Processing*, vol. 30, no. 1, pp. 32 – 46, 1985.