2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

LB-R2R-Calib: Accurate and Robust Extrinsic Calibration of Multiple Long Baseline 4D Imaging Radars for V2X

Jun Zhang[1*], Zihan Yang[1*], Fangwei Zhang[1], Zhenyu Wu[1], Guohao Peng[1], Yiyao Liu[1], Qiyang Lyu[1], Mingxing Wen[2] and Danwei Wang[1], *Fellow, IEEE*

***Abstract*— As a new sensor, 4D radar $(x, y, z, velocity)$ has great potential for V2X, due to its 3D point cloud, direct doppler velocity output, long distance ranging, low-cost, and more importantly, robust perception in all weathers. However, the extrinsic calibration of multiple long baseline 4D radars is rarely researched in V2X, which is the key to fuse multi-radars. The main reasons are three-folds: (1) New sensor. Thus, it is not surprising that little related work can be found. (2) Long baseline and large viewpoint-difference. Current works are mainly focused on unmanned vehicles, which is short baseline and small viewpoint-difference. (3) Sparse, noisy, and very cluttered 4D radar point cloud. Thus, it is challenging to rapidly and accurately locate the target and extract the feature. In this paper, *LB-R2R-Calib* (Long Baseline Radar to Radar extrinsic Calibration) is proposed to address these problems. The novelties are: (1) A new target is introduced: an eight-quadrant corner reflector enclosed by a foam sphere. The benefit is the target center is a viewpoint-invariant feature. Thus, it is ideal for large viewpoint-difference calibration. (2) A new feature extraction algorithm is proposed to rapidly locate the target and extract the target center from a very cluttered point cloud, as we observed some important characteristics of 4D radar. Experiments with two 4D radars in real environments with four configurations demonstrate our method is highly accurate and robust.**

I. Introduction

In Vehicle-to-Everything (V2X), Roadside Perception Unit (RSU) is an important component to improve the perception ability for unmanned vehicles [1]. Moreover, fusing multiple RSU sensors can further improve the perception by extending the Field-of-View (FOV) and avoiding blind spots. For example, 3D LiDAR RSU system in [2], [3], 3D radar RSU system in [4], [5], 4D radar RSU system in [6]–[8].

As a new entrant, 4D Radar $(x, y, z, velocity)$ has demonstrated great potential by combining the advantages of both 3D radar $(x, y, velocity)$ and 3D LiDAR, because it not only can provide 3D spatial information, but also can work robustly in all-weather conditions, with a low-cost. Thus, it has received more and more attention in unmanned vehicle [9], [10], [11], and V2X [6], [12]–[14]. However, as a new sensor, the extrinsic calibration among multiple RSU 4D radars has not been explored so far. This is a **new problem**, the main challenges are:

(1) As a new sensor, 4D radar just enters the market for a short period. Thus, only limited works can be found. Most current works are focused on 3D radar and short baseline calibration [15], [16]. Recently, some works are proposed for 4D radar and short baseline calibration [9], [17]–[20]. But little work can be found on 4D radar and long baseline calibration.

(2) Long baseline and large viewpoint difference makes it difficult to extract common features. Current attentions are mainly paid to sensors mounted on unmanned vehicle [9], [17], [19], [20], or a single sensor suite mounted to the road side [12], [18]. The baseline and viewpoint-difference between the sensors is short and small. The targets used for calibration are commonly planar targets [19], or single corner

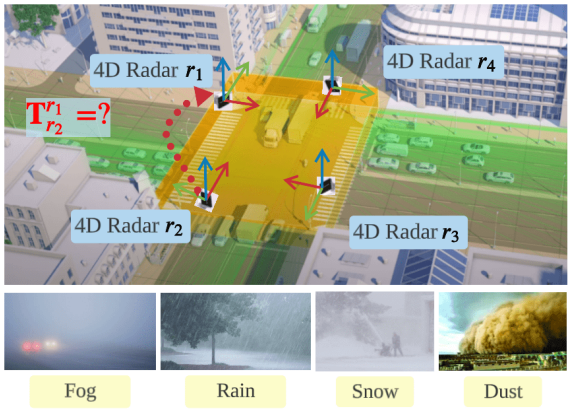


Fig. 1: *LB-R2R-Calib* is proposed to acquire the extrinsic parameters of multiple 4D radars which are installed at the road side with long baseline and large viewpoint-difference. It enables all-weather perception, expanded FOV, and reduced blind spots. As shown, four 4D Radars are installed at a junction (r_1, r_2, r_3, r_4). The frame of each radar is denoted in red $(x-)$, green $(y-)$, and blue $(z-)$ arrows. It can be observed that the baseline and viewpoint difference among the radars is long and large. The extrinsic calibration is challenging. The extrinsic parameter from r_2 to r_1 is denoted as $\mathbf{T}_{r_2}^{r_1}$. [Image source: internet[1]]

This Research is supported by China-Singapore International Joint Research Institute under Programme No.260-A022001, National Research Foundation, Singapore under its Medium Sized Centre for Advanced Robotics Technology Innovation (CARTIN), the Agency for Science, Technology and Research (A*STAR) under its National Robotics Programme (Project No. M22NBK0109), and the National Research Foundation, Singapore and Maritime and Port Authority of Singapore under its Maritime Transformation Programme (Project No. SMI-2022-MTP-04). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore and Maritime and Port Authority of Singapore.

[1] The authors are with School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore {jzhang061, yang0758, fzhang018, zhenyu002, peng0086, liuy0185, qlyu003}@e.ntu.edu.sg, edwwang@ntu.edu.sg

[2] Mingxing Wen is with Intelligent Perception Innovation Research Center, China-Singapore International Joint Research Institute. wenmx@csijri.com

* Co-first authorship and corresponding author

[1]The Future of Smart Cities – Traffic Management Sensor (smartmicro): https://youtu.be/GiOUxcJdsr4?si=UdqD-C4et-wItOuJ&t=19

reflector [9], [17], [18]. Those targets are viewpoint-sensitive, thus not friendly for large viewpoint-difference calibration. Some works use the vehicles in the traffic as features for calibration [4], [5], [21]–[23]. It is convenient and time-saving. However, the accuracy cannot reach the optimal.

(3) Sparse, noisy and very cluttered 4D radar point cloud makes it difficult to locate the target and extract the target center in complex outdoor traffic scenario. Compared with 3D LiDAR, the point cloud of 4D radar is much more sparse, noisy and cluttered [9], [24]. Thus, it is difficult to accurately, robustly and rapidly locate the target from the 4D radar.

To solve these problems, *LB-R2R-Calib* is proposed: (1) A **new target** is introduced: A eight-quadrant corner reflector which is enclosed by a foam sphere. The benefit is the target center is viewpoint-invariant, it can be detected from any viewpoints. Thus, it is especially suitable for large viewpoint-difference calibration. (2) A **novel algorithm** for target detection and feature extraction is proposed. It can swiftly pinpoint the target's location and accurately extract the target center, even in highly cluttered point cloud data. It is because we observed that the points belonging to the target can fit to a line, and the line always passes through the origin of 4D radar. Moreover, the intensity of these points will first increase then decrease.

The contributions of this paper can be summarized as:

- We propose a novel extrinsic calibration method named *LB-R2R-Calib*, which is accurate and robust for multiple roadside 4D Radars with long baseline and large viewpoint difference. As far as the authors know, this is a pioneer work.
- We introduce a new target, which can be detected from any viewpoints. The target center is viewpoint-invariant. Thus, it is friendly for long baseline and large viewpoint difference calibration.
- We propose a new strategy to rapidly locate the target and accurately extract the target center from very cluttered 4D radar point cloud, based on an important observation of the 4D radar.
- The proposed method has been validated in real-world experiments under different baseline and viewpoint difference. It demonstrates the accuracy and robustness. A benchmark is set: the 3D matching error (RMSE) of the target centers can achieve $7 \sim 25cm$ (baseline $48m$).

The structure of this paper: Sec.II surveys related works. Sec.III details the proposed method *LB-R2R-Calib*. Sec.IV presents the experiments and results. Sec.V summarizes the paper and discusses the limitation and future work.

II. Related Work

A. Short Baseline, 3D Radar related

For target-based, 3D radar-3D LiDAR calibration is performed in [16], a trihedral behind a triangular board is used as the target. A joint calibration toolbox of 3D radar-LiDAR-camera is proposed in [25], a trihedral stickered behind a four-circular-holes board is used as the target. But those targets are viewpoint-sensitive, thus are not friendly for long baseline and large viewpoint-difference calibration.

For target-less based, OpenCalib [26] is proposed for 3D radar-3D LiDAR-camera calibration, but it requires manual selection of feature points. In [27], 3D radar-rgb camera calibration is automatically performed with a neural network. Recently, the moving trajectories of vehicle or human are found to be an easy-to-obtain feature for 3D radar calibration on unmanned vehicles [23], and roadside unit [4], [5], [28]. However, those methods cannot reach the optimal accuracy. Because they rely on accurate detection of human or vehicle.

B. Short Baseline, 4D Radar Related

Current works are mainly targeted for unmanned vehicles and robots. For example, in [17], 4D radar-thermal camera calibration is performed with a spherical-trihedral target. In [18], 4D radar-rgb-LiDAR unified calibration is achieved with a novel target pair. In TJ4DRadSet [9], 4D radar-LiDAR calibration is done by manual measurement and fine-tuning with angular reflectors. In VoD [19], LiDAR-Camera-Radar joint calibration is achieved by the toolbox [29], which uses a planar board with a trihedral stickered on the back. However, those targets are not friendly for long baseline and large viewpoint-difference calibration. New target needs to be designed.

There also exist some target-less methods. For example, online calibration of 4D radar-camera is performed in [20], which is based on hand-eye calibration. In [30], the moving velocity of objects are utilized for calibration. In K-Radar [22], 4D Radar-3D LiDAR calibration is done using vehicles' bounding box as features. However, [20] and [30] require the movement of the vehicle, it is inapplicable to fixed RSU. For [22], the feature is object-level (vehicle), thus the accuracy cannot reach optimal.

C. Long Baseline, 4D Radar Related

Little literature can be found on long baseline multiple 4D radars calibration. The most related work is [21], where human movement trajectories are used for calibration of multiple 4D radar networks. However, the experiments were conducted solely within a compact indoor setting ($7 \times 4m$). It is not validated in outdoor and long baseline setting ($> 30m$) . Moreover, as it relies on accurate detection of the human, the accuracy cannot reach optimal.

Overall, little attention has been paid to extrinsic calibration of multiple long baseline 4D radars in V2X so far. Thus, we propose *LB-R2R-Calib* to fill the gap.

III. Methodology

Fig.2 depicts the workflow of *LB-R2R-Calib*. It consists of three parts: (a) Feature Extraction. (b) Feature Matching. (c) Parameter Regression. The setup of multi-4D radars and the target is shown in Fig.3.

A. Target Design

Design principle: To meet the requirements of long baseline and large viewpoint-difference calibration, the most important step is the target design:

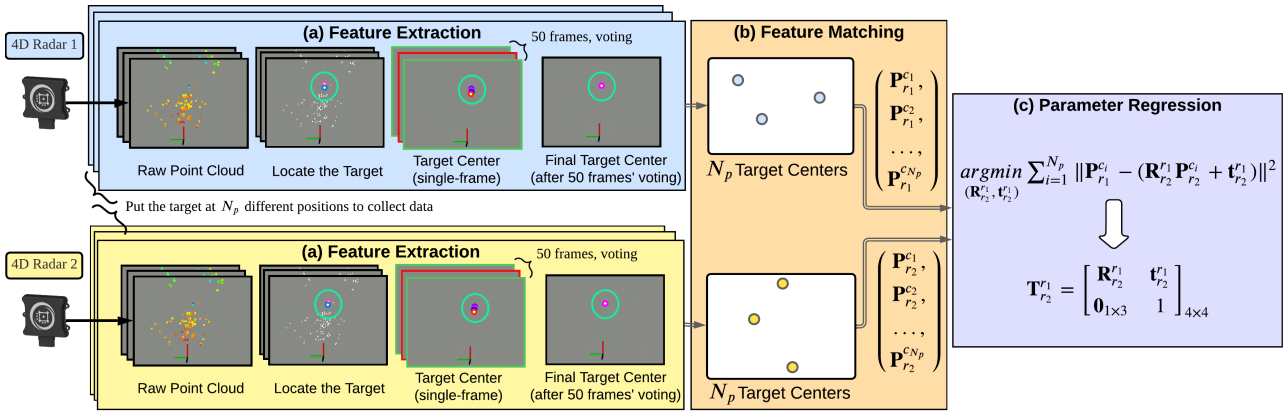


Fig. 2: Workflow of *LB-R2R-Calib*. Here, we use two 4D radars as an example. (a) ***Feature Extraction***. A new target is proposed. To rapidly locate the target and estimate the target center from a sparse, noisy and very cluttered raw point cloud, a new method is proposed based on an important observation. To avoid single-frame outliers, each radar will estimate the target center for 50 frames, then the 50 target centers will vote to output the final optimal target center. (b) ***Feature Matching***. The target will be put at N_p different positions to collect data. After data collection, each radar will get N_p target centers, denoted as $(\mathbf{P}_{r_1}^{c_1}, \mathbf{P}_{r_1}^{c_2}, ..., \mathbf{P}_{r_1}^{c_{N_p}})$ and $(\mathbf{P}_{r_2}^{c_1}, \mathbf{P}_{r_2}^{c_2}, ..., \mathbf{P}_{r_2}^{c_{N_p}})$. The correspondence is automatically built based on the target center position $c_1, c_2, ..., c_{N_p}$. (c) ***Parameter Regression***. The optimal extrinsic parameter $\mathbf{T}_{r_2}^{r_1}$ from radar 2 to radar 1 is acquired by minimizing the error in matching the 3D target centers.

1) The target should be easily detectable (or located) from a sparse, noisy and very cluttered point cloud, even from a long distance and very different viewpoints.
2) The feature point of the target, which is used for calibration, should be viewpoint-invariant, i.e., from any viewpoints, the estimated feature point should have the same spatial location.

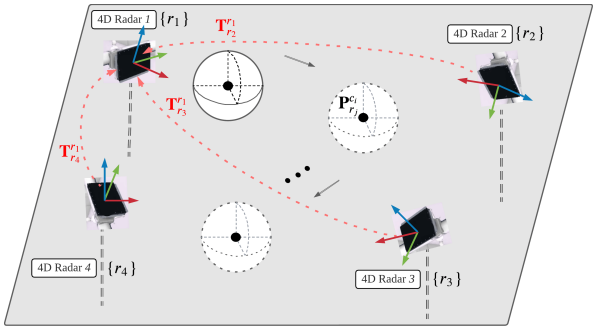


Fig. 3: Setup of four 4D radars and the target. The four radars are positioned at the same height as the traffic lights. The target is put above the ground at several positions. r_j represents the j_{th} radar $(j = 1,2,3,4)$, c_i denotes the target center at the i_{th} position $(i = 1,2,...,N_p)$. $\mathbf{P}_{r_j}^{c_i}$ denotes the target center detected by radar r_j, when the target is at i_{th} position. The extrinsic parameters from r_2, r_3, r_4 to r_1 are represented as $\mathbf{T}_{r_2}^{r_1}$, $\mathbf{T}_{r_3}^{r_1}$, $\mathbf{T}_{r_4}^{r_1}$, respectively.

Based on the above principles, a new target is designed as shown in Fig.4. The new target is inspired by [17], [31]. It is an eight-quadrant corner reflector enclosed by a foam sphere. **Advantages of the target:** (1). *Friendly for long baseline and large viewpoint difference.* The target has eight corner reflectors, therefore, it is observable from any directions. (2). *Easy-to-make and low-cost.* The corner reflectors are made by 3D printer and tin foil. The foam sphere is readily available in the market (e.g., Art Friend). (3). *General*. The target is also applicable to other sensors, such as LiDAR and camera. Since sphere target has been adopted for LiDAR and camera calibration in [2], [17], [31].

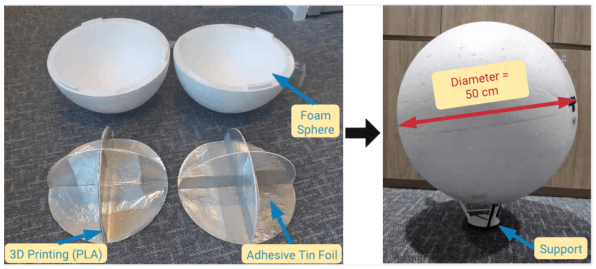


Fig. 4: Design of the new target: eight-quadrant corner reflector enclosed by a foam sphere.

B. Locate the Target

As depicted in Fig.2(a), the first step is to locate the target. It is of great importance to accurately, robustly and rapidly locate the target from very cluttered 4D radar point cloud, which is one of the key contributions of this paper. In this section, we propose a new method to achieve the objective, with a very important observation. It comprises the following steps, as shown in Fig.5:

1) Raw point cloud : The input of our algorithm is one scan of raw point cloud, as shown in Fig.5(a).

2) RCS intensity filter: When the target is put in front of the 4D radar, there will be a significant intensity difference between the target and the surrounding environment. Therefore, we use Radar Cross Section (RCS) intensity filter to filter out the points whose intensity are lower than a threshold (in our case, the threshold is 20). We just keep those high intensity points. In Fig.5(b), the white points represent the filtered out points (lower intensity), while the red points represent the kept points (higher intensity). This step helps to significantly reduce the number of points that need to be processed, facilitating the subsequent target detection.

3) Clustering and filter out single-point clusters : In this step, the high intensity points will be clustered for target detection. In Fig.5(c), the clusters are presented in various colors. Among these clusters, there exist some single-point clusters (pointed by the pink arrows), they are unlikely to be

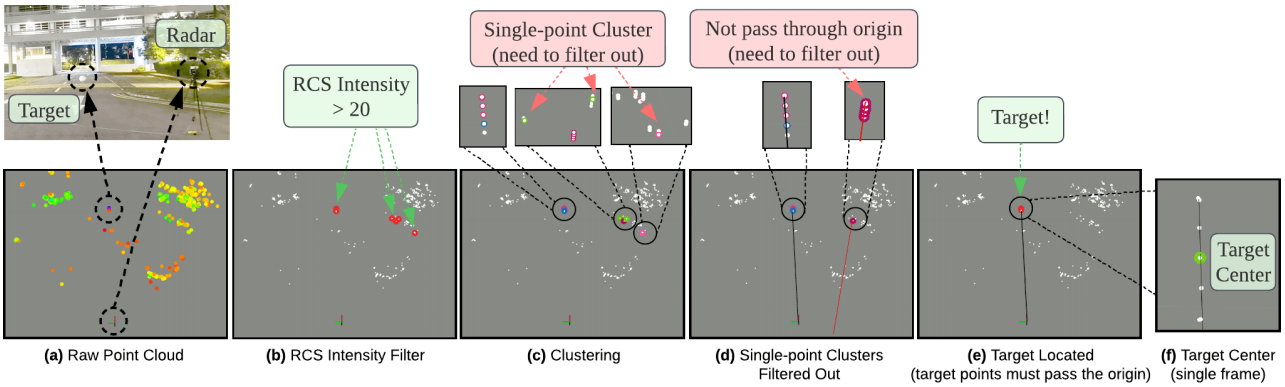


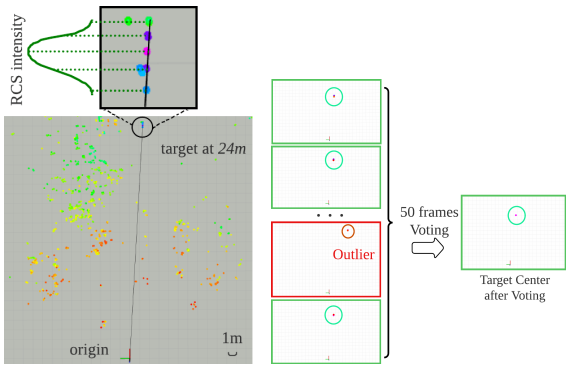
Fig. 5: The process to rapidly locate the target and estimate the target center from very cluttered 4D radar point cloud.

the target. Because from our observation, the target normally has $4 \sim 5$ points. Those single-point clusters need to be filtered out. After being filtered out, the output is shown in Fig.5(d). Now, there are only two clusters left. In the next step, we will differentiate which is the target.

4) Locate the target: For the two clusters obtained in the previous step, we need to differentiate which one is target. This is based on an important observation. We found that the points belonging to the target always satisfy three rules, as shown in Fig.6a:

- The points belonging to the target always can fit to a straight line.
- The fitted line always passes through the origin of the 4D radar frame.
- The RCS intensity of those points will first increase then decrease.

Thus, we perform RANSAC line fitting on the two clusters. From the fitted lines, we choose the one which is closest to pass the origin. In Fig.5(d), the black line passes through the origin, while the red line does not. Thus, the cluster corresponds to the black line is the target. The output result is shown in Fig.5(e). Now, the target is found!



(a) (b)

Fig. 6: (a) Our proposed method is capable to quickly locate the target. It is mainly based on this **important observation**. By zooming in the points belonging to the target, it can be observed that: those points can fit to a line (black solid line); this line always passes through the origin of radar frame; and the RCS intensity of these points will first increase then decrease (purple denotes the highest intensity). (b) The multi-frames voting scheme.

C. Target Center Estimation (Single-frame)

Now the target is located, but there are 5 points belonging to the target (Fig.5(f)). Which point is the target center? We treat the highest intensity point as the target center, because we found that the highest intensity point normally is in the middle (purple point in Fig.6a). As shown in Fig.5(f), the green point is the point with the highest intensity. Thus, it is considered as the target center for this single frame.

D. Final Target Center (Multiple-frames Voting)

To now, we can get the target center for a single frame. But, we found that the result of one single frame is not absolutely stable. Because the real environment is complex, various objects and unstable noise will cause the incorrect detection. An example is shown in Fig.6b. It shows the estimated target center of 50 frames: most estimated target centers are correct (green boundary), while one frame is outlier (red boundary).

To avoid such single-frame outliers, we adopt a simple voting strategy: for 50 target centers, we cluster them based on their distances. The cluster which has the most number of center points will be selected out as the target center cluster. The centriod of this cluster is the final optimal target center, as shown in Fig.6b.

E. Extrinsic Parameter Regression

To collect enough data, we put the target at N_p positions. Each radar will output N_p target centers $(\mathbf{P}_{r_1}^{c_1}, \mathbf{P}_{r_1}^{c_2}, ..., \mathbf{P}_{r_1}^{c_{N_p}})$ for radar 1, and $(\mathbf{P}_{r_2}^{c_1}, \mathbf{P}_{r_2}^{c_2}, ..., \mathbf{P}_{r_2}^{c_{N_p}})$ for radar 2. The correspondence between the center pairs is automatically built based on the subscript: target position $c_1, c_2, ..., c_{N_p}$.

By minimizing the matching error of the 3D target center pairs, we can obtain the optimal extrinsic parameter. The loss function is shown in Fig.2(c), the physical meaning is the optimal rotation matrix $\mathbf{R}_{r_2}^{r_1}$ and translation vector $\mathbf{t}_{r_2}^{r_1}$ should minimize the 3D matching error.

$$\underset{\mathbf{R}_{r_2}^{r_1}, \mathbf{t}_{r_2}^{r_1}}{\arg\min} \sum_{i=1}^{N_p} \| \mathbf{P}_{r_1}^{c_i} - (\mathbf{R}_{r_2}^{r_1} \mathbf{P}_{r_2}^{c_i} + \mathbf{t}_{r_2}^{r_1}) \|_2^2 \tag{1}$$

TABLE I: Calibration results and baseline for the four datasets.

Datasets	t_x (m)	t_y (m)	t_z (m)	ψ $(^\circ)$	θ $(^\circ)$	ϕ $(^\circ)$	Baseline
AAPerp	6.64	-8.28	3.74	-25.94	20.94	85.16	$11m$
AAOpp	47.29	-0.59	11.79	3.24	28.10	-177.76	$48m$
GGPerp	12.32	-23.92	0.69	0.04	0.00	89.97	$27m$
GGPara	0.37	3.22	-0.11	0.04	-0.10	0.00	$3m$

IV. Experiments

A. Experimental Configuration

We conducted experiments in real-world environments, with four setups. For each setup, two 4D radars are mounted on tripod separately and connected to a computer by Ethernet cables. The type of 4D radar is Oculii Eagle. Its detection range is $0m \sim 400m$, the azimuth and elevation FOV are $120^\circ \times 30^\circ$, the azimuth and elevation angle resolution are $0.5^\circ \times 1^\circ$. The ranging error of the 4D radar is $0.00215r$, r is the distance of a point, e.g., in our case, the target is within $50m$, so the ranging error is $0.10m$. All algorithms are running on a laptop with Ubuntu 20.04 and ROS Noetic.

The four setups are shown in the first row of Fig.8.

- **AAPerp:** Both radars are in the air (AA). The viewpoint is almost perpendicular (Perp). The baseline is around $11m$. The target is on the ground. A human will put the target at $N_p = 30$ positions for data collection. This configuration is similar to r_1 and r_2 in Fig.1.
- **AAOppo:** Both radars are in the air (AA). The viewpoint is opposite (Oppo). The baseline is around $48m$. This configuration is similar to r_1 and r_3 in Fig.1.
- **GGPerp:** Both radars are above the ground (GG) in a carpark. The viewpoint is almost perpendicular (Perp). The baseline is around $27m$.
- **GGPara:** Both radars are above the ground (GG) in a carpark. The viewpoint is almost parallel (Para). The baseline is around $3m$.

B. Quantitative Analysis

In real environment, the ground truth extrinsic parameter cannot be obtained, thus we use the 3D matching error (RMSE) of the target center pairs to quantitatively evaluate the calibration error. Our estimated extrinsic parameters are shown in Tab.I. To draw the statistics boxplot, for each setup, each time we randomly sample $N = (3,4,...,29)$ data from the 30 collected datasets. Then we use our estimated extrinsic parameter to calculate RMSE of all 30 pairs of target centers. For each N, this will be repeated for 20 times. The boxplots are shown in Fig.7. The corresponding numerical errors are shown in Tab.II. We can see that:

1) When the target positions N_p is small, the RMSE is high, and the corresponding variance is also large.
2) The RMSE and variance decreases quickly as the target positions increase. For all four datasets, the RMSE converges at $N_p = 10$ positions.
3) The final RMSE for all four datasets converges to centimeter-level. In the best-case (AAperp), the RMSE is only $0.07m$. In the worst-case (GGpara), it is $0.25m$. The errors are relatively small, considering that the baseline is long (up to $48m$), the viewpoint difference is large (up to $90^\circ, 180^\circ$) and the sensor noise is $0.10m$ for a point at $50m$.
4) The error of AAPerp is lower than AAOpp, and GGPerp is lower than GGPara. That means, if the two radars have a perpendicular viewpoint (90°), the calibration is more accurate.

C. Qualitative Analysis

To qualitatively evaluate the calibration accuracy, we performed three experiments, as demonstrated in Fig.8:

1) Registering Raw Point Cloud: The raw point clouds of the two radars are registered together, as shown in second row of Fig.8. Green point cloud belongs to radar 1. Red point cloud belongs to radar 2. It can be observed that, the registered point cloud overlaps well. To better showcase the results, we zoom in the target's point cloud, as shown in the rectangle. It can be observed that, from each radar, the target points from the two radars are overlapped well.

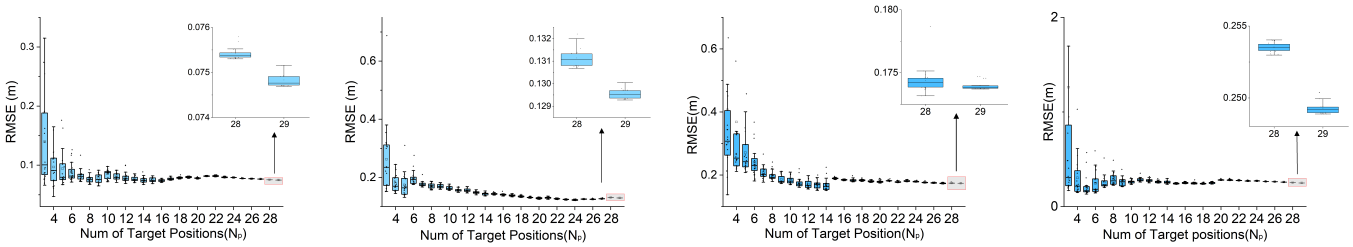
2) Registering the Target Center Pairs (Static Target): In this experiment, we register the 30 pairs of target centers. The result is shown in the third row of Fig.8. It can be observed that the target centers are nearly overlapped. Those target centers are generated by keeping the target static. In the next experiment, we will hold the target and move without a stop, then assess the target detection and target center registration results.

3) Registering a Walking Human's Point Cloud (Dynamic Target): In this experiment, we want to evaluate when the target is moving without stop, how accurate is the target detection and target center registration. To perform this, a human is walking in trajectory shape $\wedge, \infty$ in front of the two radars, with the target holding above the head. The target center is detected in real-time. The registered trajectory of the moving target is shown in the fourth and fifth row of Fig.8. Fourth row shows the top view, while fifth row shows the side view. We can observe that after registration, the trajectories of the target center overlap well, even when the target has moved for a long distance ($30m$).

Thus, we can conclude that our calibration is quite accurate in different scenarios. The corresponding video results can be found at: `https://github.com/yzh-721/LB-R2R-Calib`.

V. Conclusion and Future Work

Multi-4D radar system has great potential for V2X, intelligent transportation and unmanned vehicle, due to its all-weather perception, low-cost, extended FOV and reduced blind spots. However, the extrinsic calibration of multiple long baseline 4D radars is rarely explored. This paper introduces *LB-R2R-Calib* to address this gap: 1). A **new target** is introduced. The target center is viewpoint-invariant, thus it is suitable for large viewpoint-difference calibration. 2). A **novel strategy** is proposed to locate the target and extract the target center from sparse, noisy and very cluttered point cloud. The strategy is robust, accurate and rapid. 3) Four real experiments demonstrate *LB-R2R-Calib* is highly accurate



(a) AAPerp (b) AAOpp (c) GGPerp (d) GGPara

Fig. 7: 3D matching error of the target center pairs (RMSE) *vs.* the number of target positions N_p, for the four datasets.

TABLE II: 3D matching error of target center pairs (RMSE, unit: *m*) *vs.* No. of target positions N_p, for the four datasets.

N_p	AAPerp				AAOpp				GGPerp				GGPara			
	min	max	mean	median	min	max	mean	median	min	max	mean	median	min	max	mean	median
3	0.07	0.31	0.14	**0.10**	0.15	0.69	0.26	**0.23**	0.14	1.91	0.41	**0.31**	0.17	11.85	1.26	**0.31**
5	0.07	0.18	0.09	**0.08**	0.13	0.31	0.17	**0.16**	0.21	0.46	0.27	**0.24**	0.13	0.57	0.21	**0.17**
7	0.07	0.12	0.08	**0.08**	0.17	0.21	0.18	**0.18**	0.18	0.27	0.21	**0.20**	0.20	0.51	0.26	**0.23**
9	0.07	0.09	0.08	**0.08**	0.16	0.18	0.17	**0.17**	0.17	0.21	0.19	**0.18**	0.21	0.39	0.26	**0.26**
10	0.07	0.10	0.09	**0.09**	0.16	0.18	0.16	**0.16**	0.17	0.21	0.18	**0.18**	0.23	0.31	0.26	**0.26**
15	0.07	0.08	0.07	**0.08**	0.14	0.15	0.14	**0.14**	0.19	0.20	0.19	**0.19**	0.24	0.25	0.24	**0.24**
20	0.08	0.08	0.08	**0.08**	0.12	0.13	0.13	**0.13**	0.17	0.19	0.18	**0.18**	0.28	0.30	0.28	**0.28**
25	0.08	0.08	0.08	**0.08**	0.12	0.13	0.13	**0.13**	0.17	0.18	0.18	**0.18**	0.26	0.27	0.27	**0.26**
29	0.07	0.08	0.07	**0.07**	0.13	0.13	0.13	**0.13**	0.17	0.17	0.17	**0.17**	0.25	0.25	0.25	**0.25**

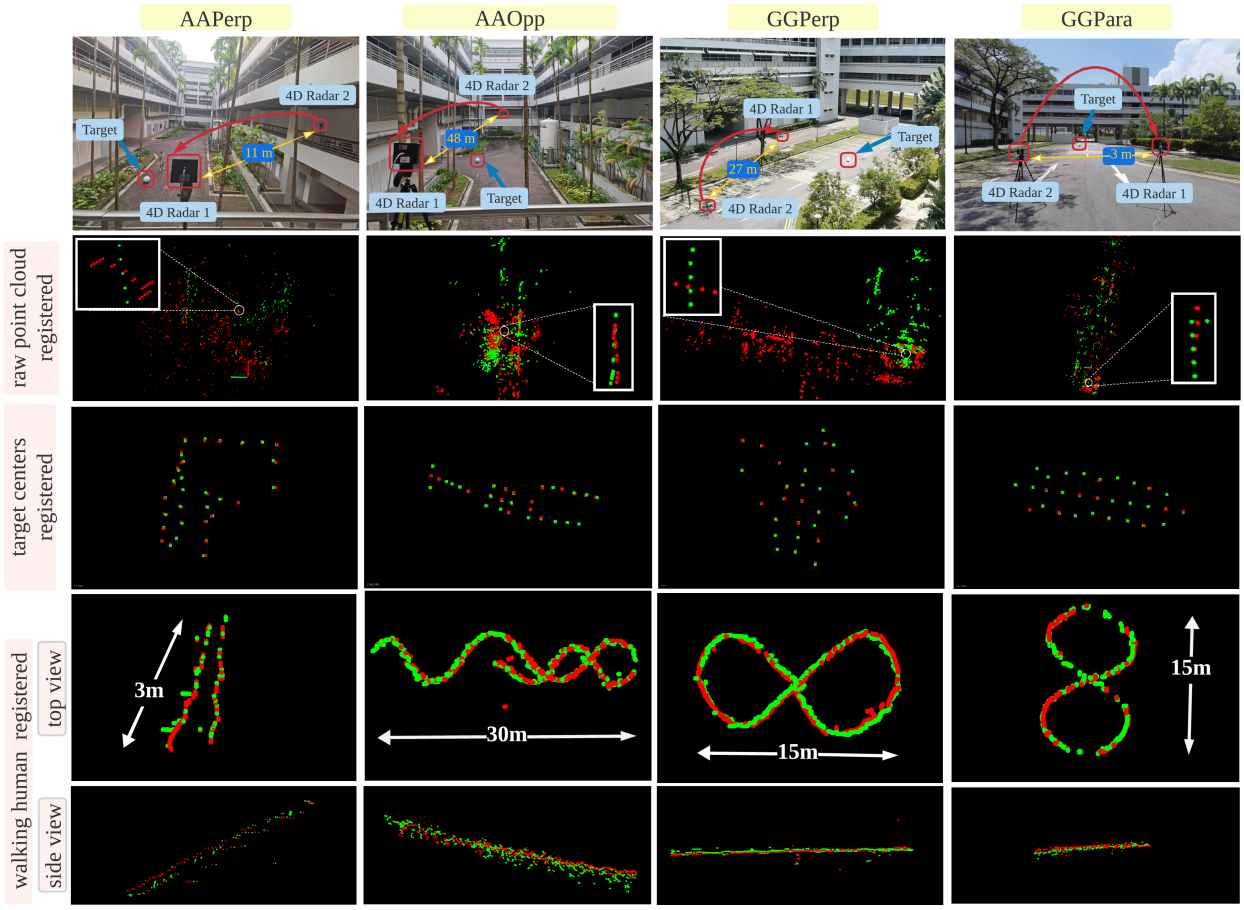


Fig. 8: Qualitative analysis by visualizing the registered point cloud with our estimated extrinsic parameter.green: radar 1, red: radar 2. *Note: As for "walking human registered", for **AAPerp, GGPerp, GGPara**, the human is walking with the target, the target center point is plotted in shape $\wedge, \infty$. It is more accurate and stable; For **AAOpp**, the human is walking without the target, the point with the maximum velocity is plotted in shape $\sim\sim$, thus, the error in the side view is larger, as the maximum velocity point on the human is not stable.

and robust: even for baseline 48*m*, opposite viewpoint, ranging noise 0.10*m*, our method still can achieve high accuracy: 3D matching error 0.13*m*. In the future, we will develop an online calibration method, without using this target.

REFERENCES

[1] H. Ren, S. Zhang, S. Li, Y. Li, X. Li, J. Ji, Y. Zhang, and Y. Zhang, "TrajMatch: Toward Automatic Spatio-Temporal Calibration for Roadside LiDARs Through Trajectory Matching," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2023.

[2] J. Zhang, Q. Lyu, G. Peng, Z. Wu, Q. Yan, and D. Wang, "LB-L2L-Calib: Accurate and Robust Extrinsic Calibration for Multiple 3D LiDARs with Long Baseline and Large Viewpoint Difference," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 926–932.

[3] J. Zhang, Q. Yan, M. Wen, Q. Lyu, G. Peng, Z. Wu, and D. Wang, "LB-L2L-Calib 2.0: A Novel Online Extrinsic Calibration Method for Multiple Long Baseline 3D LiDARs using Objects," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023.

[4] J. Bai, S. Li, H. Zhang, L. Huang, and P. Wang, "Robust Target Detection and Tracking Algorithm Based on Roadside Radar and Camera," *Sensors*, vol. 21, no. 4, 2021.

[5] C. Zhang, J. Wei, J. Dai, S. Qu, X. She, and Z. Wang, "A Roadside Millimeter-Wave Radar Calibration Method Based on Connected Vehicle Technology," *IEEE Intelligent Transportation Systems Magazine*, vol. 15, no. 3, pp. 117–131, 2023.

[6] M. Lei, D. Yang, and X. Weng, "Integrated Sensor Fusion Based on 4D MIMO Radar and Camera: A Solution for Connected Vehicle Applications," *IEEE Vehicular Technology Magazine*, vol. 17, no. 4, pp. 38–46, 2022.

[7] S. Yao, R. Guan, Z. Peng, C. Xu, Y. Shi, Y. Yue, E. G. Lim, H. Seo, K. L. Man, X. Zhu, and Y. Yue, "Radar Perception in Autonomous Driving: Exploring Different Data Representations," *arXiv preprint arXiv:2312.04861*, 2023.

[8] S. Yao, R. Guan, X. Huang, Z. Li, X. Sha, Y. Yue, E. G. Lim, H. Seo, K. L. Man, X. Zhu, and Y. Yue, "Radar-Camera Fusion for Object Detection and Semantic Segmentation in Autonomous Driving: A Comprehensive Review," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 2094–2128, 2024.

[9] L. Zheng, Z. Ma, X. Zhu, B. Tan, S. Li, K. Long, W. Sun, S. Chen, L. Zhang, M. Wan, L. Huang, and J. Bai, "TJ4DRadSet: A 4D Radar Dataset for Autonomous Driving," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 2022, pp. 493–498.

[10] J. Zhang, H. Zhuge, Z. Wu, G. Peng, M. Wen, Y. Liu, and D. Wang, "4DRadarSLAM: A 4D Imaging Radar SLAM System for Large-scale Environments based on Pose Graph Optimization," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 8333–8340.

[11] B. Tan, Z. Ma, X. Zhu, S. Li, L. Zheng, S. Chen, L. Huang, and J. Bai, "3-D Object Detection for Multiframe 4-D Automotive Millimeter-Wave Radar Point Cloud," *IEEE Sensors Journal*, vol. 23, no. 11, pp. 11 125–11 138, 2023.

[12] S. Agrawal, R. Song, K. Doycheva, A. Knoll, and G. Elger, "Intelligent Roadside Infrastructure for Connected Mobility," in *Smart Cities, Green Technologies, and Intelligent Transport Systems*, C. Klein, M. Jarke, J. Ploeg, M. Helfert, K. Berns, and O. Gusikhin, Eds. Cham: Springer Nature Switzerland, 2023, pp. 134–157.

[13] J. Niu, Y. F. Wu, Y. J. Cheng, L. W. Mou, T. J. Li, H. N. Yang, S. L. Pu, L. J. Shen, L. Shen, and T. Qian, "A Sparse Array for 77 GHz 4D High-Resolution Imaging Radar Based on Entropy Model and Convex Optimization," in *2021 IEEE International Workshop on Electromagnetics: Applications and Student Innovation Competition (iWEM)*, vol. volume1, 2021, pp. 1–3.

[14] Z. Han, J. Wang, Z. Xu, S. Yang, L. He, S. Xu, and J. Wang, "4D Millimeter-Wave Radar in Autonomous Driving: A Survey," *arXiv preprint arXiv:2306.04242*, 2023.

[15] P. Jiang and S. Saripalli, "Improving Extrinsics between RADAR and LIDAR using Learning," in *2023 IEEE Intelligent Vehicles Symposium (IV)*, 2023, pp. 1–6.

[16] C.-L. Lee, Y.-H. Hsueh, C.-C. Wang, and W.-C. Lin, "Extrinsic and Temporal Calibration of Automotive Radar and 3D LiDAR," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020.

[17] J. Zhang, S. Zhang, G. Peng, H. Zhang, and D. Wang, "3DRadar2ThermalCalib: Accurate Extrinsic Calibration between a 3D mmWave Radar and a Thermal Camera Using a Spherical-Trihedral," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 2022, pp. 2744–2749.

[18] S. Agrawal, S. Bhanderi, K. Doycheva, and G. Elger, "Static multi-target-based auto-calibration of RGB cameras, 3D Radar, and 3D Lidar sensors," *IEEE Sensors Journal*, pp. 1–1, 2023.

[19] A. Palffy, E. Pool, S. Baratam, J. F. P. Kooij, and D. M. Gavrila, "Multi-Class Road User Detection With 3+1D Radar in the View-of-Delft Dataset," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4961–4968, 2022.

[20] E. Wise, J. Peršić, C. Grebe, I. Petrović, and J. Kelly, "A Continuous-Time Approach for 3D Radar-to-Camera Extrinsic Calibration," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 13 164–13 170.

[21] A. Shastri, M. Canil, J. Pegoraro, P. Casari, and M. Rossi, "mmSCALE: Self-Calibration of mmWave Radar Networks from Human Movement Trajectories," in *2022 IEEE Radar Conference (RadarConf22)*, 2022, pp. 1–6.

[22] D.-H. Paek, S.-H. Kong, and K. T. Wijaya, "K-Radar: 4D Radar Object Detection for Autonomous Driving in Various Weather Conditions," in *Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2022.

[23] W. Zhangu, Z. Jun, D. Chuanguang, G. Xin, and Y. Kai, "Traffic Vehicle Cognition in Severe Weather Based on Radar and Infrared Thermal Camera Fusion," *Measurement Science and Technology*, vol. 32, no. 9, p. 095111, jun 2021.

[24] J. Zhang, H. Zhuge, Y. Liu, and et al., "NTU4DRadLM: 4D Radar-centric Multi-Modal Dataset for Localization and Mapping," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, 2023.

[25] J. Domhof, J. F. Kooij, and D. M. Gavrila, "An Extrinsic Calibration Tool for Radar, Camera and Lidar," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 8107–8113.

[26] G. Yan, Z. Liu, C. Wang, C. Shi, P. Wei, X. Cai, T. Ma, Z. Liu, Z. Zhong, Y. Liu, M. Zhao, Z. Ma, and Y. Li, "OpenCalib: A Multi-sensor Calibration Toolbox for Autonomous Driving," *Software Impacts*, vol. 14, p. 100393, 2022.

[27] C. Schöller, M. Schnettler, A. Krämmer, G. Hinz, M. Bakovic, M. Güzet, and A. Knoll, "Targetless Rotational Auto-Calibration of Radar and Camera for Intelligent Transportation Systems," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 3934–3941.

[28] C. Zhang, J. Wei, A. S. Hu, and P. Fu, "A Novel Method for Calibration and Verification of Roadside Millimetre-wave Radar," *IET Intelligent Transport Systems*, vol. 16, no. 3, pp. 408–419, 2022.

[29] J. Domhof, J. F. P. Kooij, and D. M. Gavrila, "A Joint Extrinsic Calibration Tool for Radar, Camera and Lidar," *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 3, pp. 571–582, 2021.

[30] Y. Bao, T. Mahler, A. Pieper, A. Schreiber, and M. Schulze, "Motion Based Online Calibration for 4D Imaging Radar in Autonomous Driving Applications," in *2020 German Microwave Conference (GeMiC)*, 2020, pp. 108–111.

[31] J. V. Kümmerle, "Multimodal Sensor Calibration with a Spherical Calibration Target," Ph.D. dissertation, Karlsruher Institut für Technologie (KIT), 2020.