Robust Metric Localization in Autonomous Driving via Doppler Compensation With Single-Chip Radar

Robust Metric Localization in Autonomous Driving via Doppler Compensation With Single-Chip Radar

Pengen Gao⬥, Shengkai Zhang⬥, *Member, IEEE*, Wei Wang⬥, *Senior Member, IEEE*, and Chris Xiaoxuan Lu⬥

Abstract— Metric localization is vital to autonomous driving where it corrects cumulative errors in a long-term run. Such errors are inevitable in real scenarios where GPS signals or some other drift-free exteroceptive measurements are not available, *e.g.*, when an automobile goes through a tunnel. Using FMCW-based mmWave radars is an attractive metric localization technique with improved robustness as RF signals can traverse small particles in harsh weather conditions like snowing, foggy, and storming, but it faces a fundamental challenge of Doppler distortion. Existing works take spatial constraints to mitigate the Doppler distortion of point clouds from mechanical radars with limited accuracy. Modern single-chip mmWave radars that provide dynamic estimates, *i.e.*, radial velocities, bring new opportunities to develop more accurate approaches. This paper presents `DC-Loc++`, a robust metric localization framework by compensating Doppler distortions using a single-chip mmWave radar. It consists of an explicit velocity-assisted Doppler compensation module for each radar sub-map, an uncertainty-aware metric registration algorithm, and a failure recovery method that validates measurement constraints to generate a more confident pose graph for optimizing vehicle poses. Extensive experiments on both nuScenes dataset and a synthetic CARLA dataset show the effectiveness of `DC-Loc++`, achieving 99.2% success rate and more than 20.0%, 30.2% error reductions in terms of translation and rotation estimates, respectively, compared with existing approaches.

Index Terms— Automotive radar, metric localization, Doppler effect.

Fig. 1. Metric localization with automotive radar: a vehicle maps its pose to a pre-built map by previous radar frames. The transformation between the current frame (highlighted in blue) and the reference radar frame (highlighted in red) can be derived from radar sub-map registration (top row) without considering illumination conditions. The middle and bottom rows show the same location taken by cameras in the daytime and night. Such a process can effectively reduce cumulative errors.

I. Introduction

METRIC localization maps a vehicle's position to one in a pre-built map based on its exteroceptive observations. It plays a vital role in obtaining accurate and

drift-free positions for autonomous driving [1]. The mainstream autonomous driving solutions use optical sensors, *e.g.*, camera and LiDAR [1], [2], to observe a vehicle's surroundings with their rich and reliable sensing features. However, the nanometre-scale wavelength of optical signals makes them unable to traverse small particles like smoke, fog, raindrops, and snowflakes, preventing them from seeing obstacles effectively in adverse weather. Complementary to optical signals, millimeter-wave (mmWave) radar signals can easily traverse through small particles and sense solid objects behind. Thus, mmWave radar based metric localization (c.f. Fig. 1) has attracted strong interests in both industry and academia [3], [4], [5], [6].

Doppler distortion is one of the fundamental challenges for mmWave radar-based approaches. mmWave radar emits Frequency-Modulation Continuous Wave (FMCW) signals

(c.f. Section III-B) from an antenna array to measure ranges and angles to objects which bounce back the signals. A slight frequency shift caused by the Doppler effect will result in a non-negligible drift of ranging measurements [7], distorting the generated point cloud. Such a distortion is dominated by the relative velocity between the radar and its target. It will result in a large mismatch of the locations where a vehicle goes through at different speeds, failing to associate a sub-map to the position. Existing works [8], [9] focus on mechanical spinning radars that take spatial constraints to infer the Doppler shift so as to compensate for the map distortion coarsely. They have inferior accuracy due to the lack of radial velocity measurements, making the impact of Doppler distortion still exist in data association, as shown in our study in [10].

Recently, compact single-chip mmWave radars (*a.k.a.* automotive radar) have been widely deployed in automobiles [11]. Although the field of view is limited to a certain degree through beamforming, it enables the estimation of radial velocities associated with spatial points. The radial velocity is more related to the Doppler effect, bringing new opportunities for developing more accurate approaches. However, as shown in our study in Section III, single-chip mmWave radars are not immune to Doppler distortion and have large metric localization errors. In this paper, we propose `DC-Loc++`, a single-chip radar metric localization framework that can robustly estimate the vehicle's pose by addressing the Doppler distortion, radar measurement uncertainties, and localization failures. Specifically, `DC-Loc++` exploits the radial velocity estimates from radar signals to restore the range shift from the Doppler effect. Such a restored radar sub-map can significantly boost localization accuracy, especially when a vehicle visits the same place at different speeds. In addition, we propose an uncertainty-aware registration algorithm to align the live radar sub-map with a pre-built one. In particular, `DC-Loc++` tracks the range, angle, and Doppler uncertainties to formulate the information matrix, which dynamically adjusts the weight of each point correspondence to yield more accurate results. Furthermore, we devise a pose graph truncating scheme to filter out metric localization failures caused by outlier radar frames. Such outliers are inevitably produced due to the specular reflection and multipath of RF signals. This module substantially improves the robustness of `DC-Loc++`.

To evaluate the performance of `DC-Loc++`, we conduct extensive experiments on both nuScenes [11] and synthetic CARLA [12] dataset. The results demonstrate that `DC-Loc++` outperforms existing approaches by reducing more than 20% and 30.2% errors in terms of translation and rotation, respectively.

Our contributions can be summarized as follows.

- We elaborate on the adverse Doppler effect in the FMCW ranging process and propose an efficient compensation method to correct the Doppler distortion in each radar sub-map for metric localization of single-chip mmWave radars.
- We present an uncertainty-aware registration algorithm that tracks measurement uncertainties to obtain more accurate pose estimates.

- We devise a pose graph truncating scheme that detects and filters out problematic measurement constraints to improve the robustness of `DC-Loc++`.
- We conduct extensive experiments on both real-world and synthetic datasets to demonstrate the effectiveness of `DC-Loc++`.

The rest of the paper is organized as follows. Section II reviews the related works on metric localization for mmWave radars. The design of `DC-Loc++` is presented in Section III. Section IV reports the experimental results of `DC-Loc++` over two datasets. Section V concludes the paper.

II. RELATED WORK

The mainstream research relies on optical sensors such as cameras and LiDARs for metric localization. However, their performance heavily degenerates under challenging environments, *e.g.*, foggy or rainy weather. On the contrary, mmWave radars have robust sensing performance with the use of large wavelengths. Moreover, the recent advances in FMCW radar technology significantly improve the sensing accuracy and range, making the radar a promising sensor for metric localization.

A. Optical Sensor Based Metric Localization

Camera is a primary sensor for vehicles' metric localization due to its low cost and general use. Accurate localization results can be obtained by associating camera images with already built LiDAR maps [13] or existing street views [1]. Some researchers also explore the use of semantic information [14] of landmarks and incorporate multi-level features [15] to deal with challenging environments. However, such methods are still sensitive to illumination changes and textureless environments.

LiDAR, which emits modulated lasers for perception, also attracts much attention in academia and industry. Existing LiDAR systems rely on geometrical constraints of features in point cloud for motion estimation and metric localization [16], [17]. Meanwhile, learning-based methods are also well studied and achieve decent performance in ideal conditions [18], [19]. These systems, however, are vulnerable to dust or smoke. They perform poorly in adverse weather conditions.

B. Radar-Based Metric Localization

RF signals can penetrate, reflect, or diffract from obstacles, making mmWave radars robust to visual degradation. Meanwhile, the signal's large wavelength poses many fundamental challenges of mapping and localization, *e.g.*, multipath reflections and low-resolution perceptions. Prior works have made great efforts to tackle such problems. Ward et al. [20] proposed a two-stage pipeline for metric localization using radars. They adopt an EKF-based framework that integrates the Iterative Closest Point (ICP) [21] algorithm with an existing map. RadarSLAM [22] demonstrated the first radar-based Simultaneous Localization and Mapping (SLAM) system that works well in adverse weather, utilizing vision-based schemes for feature extraction and graph matching. Barnes et al. [3]
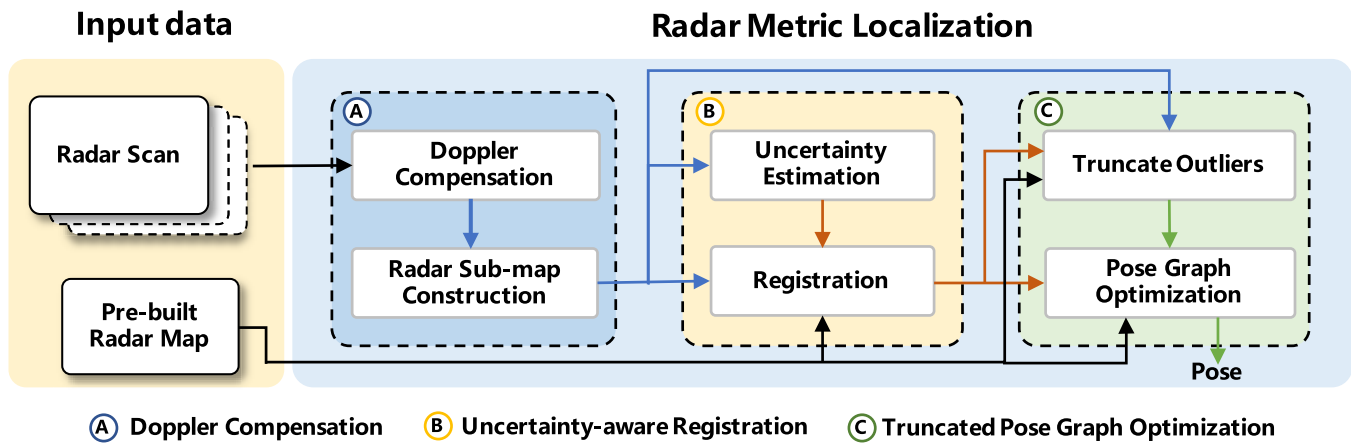
Fig. 2. The system overview of `DC-Loc++`. Given a pre-built radar map, `DC-Loc++` takes multiple radar scans from a single-chip mmWave radar installed on a vehicle to feed into the Doppler Compensation module. Then, `DC-Loc++` aggregates those compensated radar scans into a sub-map, which is aligned with the pre-built map by adopting an uncertainty-aware registration algorithm. Finally, the truncated pose graph optimization module is used to obtain robust and accurate vehicle poses.

presented an end-to-end radar localization pipeline, which predicts features along with their descriptors from raw radar inputs before feeding into a differentiable Umeyama-style estimator. Daniele et al. [6] proposed a coarse-to-fine radar localization algorithm that consolidates the place recognition and radar pose estimation algorithms into a hierarchical localization process.

Besides developing the pure radar-based sensing, more efforts explore cross-modality solutions. Yin et al. [23] combined the radar and LiDAR measurements and adopted a brute force based learning algorithm to localize the vehicles by maximizing the likelihood between the live radar scans and an existing LiDAR map. Tang et al. [4] proposed an unsupervised-learning framework that aligns 2D radar scans with the geometric structure in satellite images. Although prior works improved the metric localization performance, they neglect the impact of the Doppler effect that severely exacerbates metric localization errors, especially at high speeds. Moreover, their methods are vulnerable to large driving environment changes, especially in the presence of occlusion and dynamic objects, which pose difficulties in data association and degrade their localization accuracy.

The closest work to ours is [8]. It mitigated the Doppler distortion in the metric localization using spinning radars. Their Doppler compensation scheme has been the state-of-the-art (SOTA) solution for radar-based metric localization. However, due to the limitation of spinning radar, their method inferred the velocity from the pose estimations, which have been corrupted by the Doppler shift. In practice, we find that their metric localization error still grows larger with the increase of velocities between radar scans.

In this paper, our work focuses on explicitly eliminating the Doppler effect in metric localization using single-chip automotive radars. It compensates for the Doppler distortion of radar sub-maps and better deals with the noise floor of the measurements from low-cost automotive radars by uncertainty estimation and pose graph optimization.

III. RADAR METRIC LOCALIZATION WITH DOPPLER COMPENSATION

A. System Overview

In this work, we propose a metric localization system using a single-chip mmWave radar, termed DC-Loc++. We address the Doppler distortion and localization failures that highly degrade the radar localization performance by three modules (as shown in Fig. 2).

- **Doppler compensation module** corrects the Doppler distortion in each radar scan and creates a dense radar sub-map to facilitate the registration.
- **Uncertainty-aware registration module** incorporates ranges, angles, and radial velocities obtained from radar scans to combat the high noise level of measurements for better registrations.
- **Truncated pose graph optimization module** detects and corrects localization failures to enhance the system's robustness.

In what follows, we first elaborate on the metric localization issues caused by the Doppler distortion in Sec. III-B and then dive into the detail of our system design.

B. Doppler Shift in FMCW Ranging

1) Principles for FMCW Radar Measurement: Automotive mmWave radars usually adopt FMCW signals for measuring the range (*i.e.*, the point cloud in a scan) and relative radial velocity between the vehicle and a target object. Here we briefly introduce the principle of FMCW radar measurement [24].

When the radar receives a reflected signal from a target, it performs a dechirp operation by mixing the received signal with the transmitted signal, which produces an *Intermediate Frequency* (IF) signal $s(t)$, *i.e.*,

$$s(t) = A e^{j(2\pi f_b t + \phi)}, \tag{1}$$

where A denotes the signal attenuation, f_b and ϕ are the frequency and the phase corresponding to the central frequency
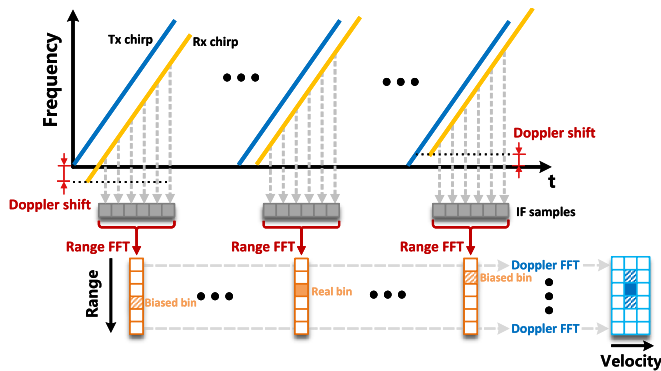
Fig. 3. The process of FMCW-based range and velocity measurements. The Doppler frequency shift occurs when the radar and objects in the environment have relative velocities. The frequency shift transforms into a biased range measurement (shaded square) for each radar target.

of $s(t)$, respectively. The range r and radial velocity v between the radar and its target can be calculated as:

$$r = \frac{cf_b}{2K}, \quad v_r = \frac{\lambda f_v}{2T_c} \quad (K = B/T_r), \tag{2}$$

where c is the speed of light, B, T_r, and λ denote bandwidth, signal duration, and signal wavelength. f_v and T_c denote the phase change rate and the time separation between IF signals.

2) Doppler Effect in Radar Measurement: The above ranging operation of FMCW radars works well when a vehicle's ego-velocity is low. However, for autonomous driving scenarios where the ego-velocity changes fast from time to time, the Doppler effect dominated by the radial velocity will impose a Doppler frequency shift on the IF signals. As a result, it will incur a non-negligible knock-on effect on the range measurements. As shown in Fig. 3, the Doppler effect causes an apparent frequency shift on the FMCW signal, eventually resulting in the Doppler distortion to the point cloud from a radar scan.

The Doppler distortion has varying impacts on different radar-based localization tasks. Specifically, radar odometry adopts successive radar scans for ego-pose estimation, and the Doppler effect typically has minor impacts on radar odometry. This is because vehicles usually travel at consistent speeds within short intervals. Therefore, the Doppler effect will impose similar impacts on successive radar scans. The spatial relationships between radar targets in successive radar scans will not change dramatically. As a result, the radar scan registration procedure in radar odometry is resilient to the Doppler effect [8]. However, the Doppler distortion will significantly degrade the performance of metric localization. This is due to the fact that metric localization aims to align live radar scans to a pre-built map, while vehicles usually revisit places at different speeds. Thus the Doppler effect will incur inconsistent distortions to mapping results, which introduces difficulties in creating accurate pose estimation.

In the FMCW radar measurement model, the Doppler frequency shift can be modeled as $f_d = \frac{2v_r}{\lambda}$ [7], where v_r denotes the radial component of the relative velocity and λ denotes the signal wavelength. Thus, given the measured IF signal frequency $\hat{f}_b$, the real IF signal frequency f_b can be

formulated as [8]:

$$f_b = \hat{f}_b - f_d, \tag{3}$$

which means a range shift r_d has been imposed to the real range r. Substituting (3) into (2), we can get:

$$r = \hat{r} - \frac{cf_d}{2K}, \tag{4}$$

where $\hat{r}$ is the radar range measurement, K represents the ramp rate of the chirp. We can observe that the latter term of (4) is the very range shift caused by the Doppler effect, *i.e.*,

$$r_d = \frac{cf_d}{2K} = \frac{f_c}{K}v_r, \tag{5}$$

where f_c denotes the frequency of the transmitted chirp signal. It is noted that the range shift r_d is proportional to the radial velocity v_r, indicating that approaching a target in environments shortens the range and vice versa. Although r_d is also inversely proportional to K, mmWave radars typically keep a low ramp rate to guarantee its range resolution [24], making velocity v_r the dominant factor distorts a radar's range scan.

Although Burnett et al. [8] also discuss the Doppler effect in radar measurements, their algorithm is designed for spinning mmWave radars. Since spinning radars cannot measure the radial velocity due to its progressive data collection mechanism, the velocity is derived from the vehicle's ego-motion, whose accuracy has been affected by the Doppler distortion.

C. Doppler Compensation in Radar Sub-Maps

The key idea behind our Doppler compensation method is that the radial velocity measured by automotive mmWave radars is resilient to the Doppler frequency shift. Recent work [7] concludes that the Doppler effect can only shift the frequency peaks in the range bin rather than their phases. And the phases are essential to estimate the radial velocity. Specifically, as shown in Fig. 3, the Doppler FFT utilizes the phases of the range bins to estimate the velocity. The FFT outputs provide the phase change rate of the range bins, having no relationship with the frequency of the IF signals. Thus, the Doppler effect has no impact on the velocity measurements, indicating that we can correct the Doppler distortion in each radar scan by using the radial velocity of each target.

1) Doppler Compensation: Specifically, let $\hat{\mathbf{p}} \in \mathbb{R}^3$ be a radar target in Cartesian coordinates, where $(\hat{\cdot})$ denotes the measured data. Given the radar measurements $\hat{\mathbf{m}} = (\hat{r}, \hat{v}_r, \hat{\phi}, \hat{\theta})$, where each variable represents for the range, radial velocity, azimuth angle, and elevation angle of the radar target, the Doppler compensated radar target $\hat{\mathbf{p}}^c$ can be expressed as:

$$\hat{\mathbf{p}}^c = \left(\hat{r} - \frac{f_c}{K}\hat{v}_r\right)[\cos\hat{\theta}\cos\hat{\phi}, \cos\hat{\theta}\sin\hat{\phi}, \sin\hat{\theta}]^\mathrm{T}. \tag{6}$$

Therefore, the Doppler-compensated radar scan at time t can be obtained by compensating the Doppler shift for each radar target, *i.e.*, $\hat{\mathcal{P}}_t^c = \{\hat{\mathbf{p}}_1^c, \hat{\mathbf{p}}_2^c, \cdots, \hat{\mathbf{p}}_n^c\}$, where n is the target number in the radar scan. Fig. 4 shows such a compensation process and illustrates the range shift for each radar target.
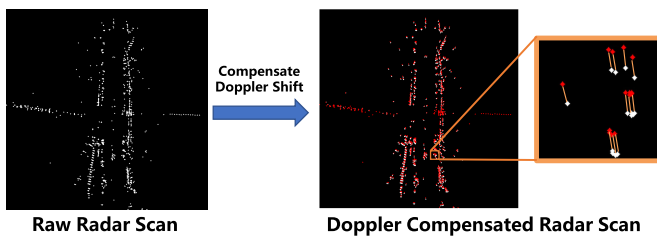
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

GAO et al.: ROBUST METRIC LOCALIZATION IN AUTONOMOUS DRIVING VIA DOPPLER COMPENSATION

5

Fig. 4. This diagram shows the Doppler compensation operation with a zoomed-in view. The Doppler compensation is performed in each radar scan to restore the Doppler distortion. The radar scan after compensation is highlighted in red. The orange lines represent the range shift r_d.
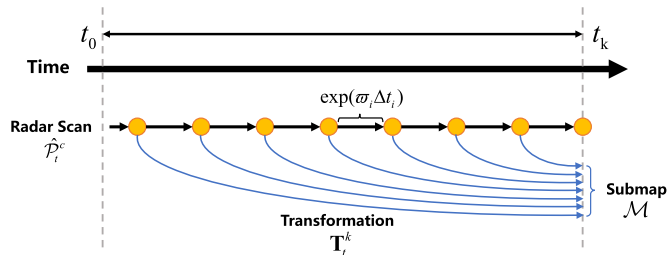

Fig. 5. This diagram shows the radar sub-map construction pipeline. We can obtain a dense radar sub-map $\mathcal{M}$ by aggregating radar scans $\hat{\mathcal{P}}_i^c$ between t_0 and t_k. The transformation matrix $\mathbf{T}_t^k$ can be obtained by integrating instantaneous ego-velocities ϖ_i in a short period Δt_i.

2) Radar Sub-Map Construction: Due to the sparsity of the point cloud produced by a mmWave radar, it is difficult to provide sufficient radar targets for metric localization with a single radar scan. Specifically, in order to get the relative pose between two radar scans, a data association process is required to match the repetitive targets in both radar scans. However, the sparsity of radar measurement deteriorates this process, making it challenging to attain accurate pose estimations of the vehicle. To address this issue, we concatenate k consecutive Doppler-compensated radar scans to get a denser radar sub-map $\mathcal{M}$ by transferring all radar scans sampled between t_0 to t_k into the same coordinate (*c.f.*, Fig. 5):

$$\mathcal{M} = \bigcup_{t=0}^{k-1} \mathbf{T}_t^k \cdot \hat{\mathcal{P}}_t^c, \qquad (7)$$

where $\mathbf{T}_t^k \in SE(3)$ represents the relative transformation from radar frame at time t to radar frame at time t_k. We denote such a superposition of point clouds as a Doppler-free radar sub-map, which is more informative for registration.

In order to get $\mathbf{T}_t^k$, we adopt a progressive scheme by integrating radar instantaneous ego-velocity[1] ϖ introduced in [25] as follows:

$$\mathbf{T}_t^k = \left(\prod_{i=t}^{k-1} \exp\left(\varpi_i \Delta t_i\right) \right)^{-1}, \quad t = \{0, 1, \cdots k-1\} \quad (8)$$

where Δt_i represents the time interval between i^{th} and $i+1^{\text{th}}$ radar scans. It is noted that when $t = 0$, $\mathbf{T}_0^k$ is the transformation between two radar keyframes, which is regarded as a coarse prediction of the localization result as discussed in Sec. III-E.2.

[1]We use the velocity profile (radial velocity over azimuth angles) of the radar targets to estimate the ego-velocity of the vehicle. Please refer to [25] for more details.
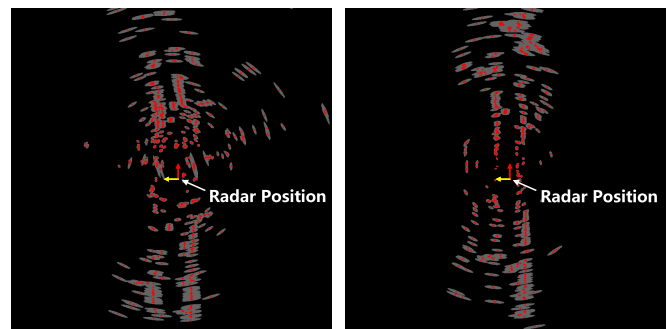

Fig. 6. The uncertainty estimation for each point (red) with their 5 standard deviation uncertainty ellipse (gray). The points further away from the radar (image center) have larger uncertainty ellipses.

D. Uncertainty-Aware Registration

Although the Doppler distortion has been corrected in the above compensation method, the measured velocities and ranges still have significant uncertainties due to low-quality radar measurements, especially when targets are far away from the radar. These low-quality measurements significantly reduce the accuracy of metric localization. Uncertainty-aware registration module addresses this problem by taking the uncertainty into an optimization framework to find the best alignment between the live radar sub-map and a pre-built radar map with minimum matching residuals.

1) Uncertainty Estimation: We first convert the radar's raw measurements from spherical coordinates to Cartesian coordinates for ease of problem formulation. Every point $\hat{\mathbf{p}}_k^c$ in radar sub-map $\mathcal{M}$ can be formulated as a function of radar measurements $\mathbf{m}_k = \left[r_k, v_{r,k}, \phi_k, \theta_k\right]^{\text{T}}$, in addition with additive Gaussian noise $\mathbf{n}_k = \left[n_r, n_{v_r}, n_\theta, n_\phi\right]^{\text{T}}$

$$\begin{aligned} \hat{\mathbf{p}}_k^c &= \mathbf{f}(\mathbf{m}_k, \mathbf{n}_k) \\ &\approx \mathbf{f}(\mathbf{m}_k) + \mathbf{J}\mathbf{n}_k, \quad \mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, \Sigma), \end{aligned} \qquad (9)$$

where $\mathbf{f}$ is the coordinate transformation function. $\Sigma = diag\left(\sigma_r^2, \sigma_{v_r}^2, \sigma_\theta^2, \sigma_\phi^2\right)$ is the covariance of radar measurement $\hat{\mathbf{m}}_k$ in spherical coordinates. In Equation (9), linearization is conducted to approximate the non-linear function, and $\mathbf{J}$ is the Jacobian matrix of the function $\mathbf{f}$ *w.r.t.* the radar measurements. The detailed expression of $\mathbf{J}$ is demonstrated as follows:

$$\begin{aligned} \mathbf{J} &= \frac{\partial \mathbf{f}(\mathbf{m}, \mathbf{0})}{\partial \left[r, v_r, \theta, \phi\right]^{\text{T}}} \\ &= \mathbf{R} \begin{bmatrix} c\theta c\phi & -\beta \cdot c\theta c\phi & -r_c \cdot s\theta c\phi & -r_c \cdot c\theta s\phi \\ c\theta s\phi & -\beta \cdot c\theta s\phi & -r_c \cdot s\theta s\phi & r_c \cdot c\theta c\phi. \\ s\theta & -\beta \cdot s\theta & r_c \cdot c\theta & 0 \end{bmatrix} \end{aligned} \qquad (10)$$

where c and s represent for $\cos(\cdot)$ and $\sin(\cdot)$ for brevity. $\beta = \frac{f_c}{K}$ is the Doppler compensation coefficient and $r_c = r - \beta v_r$ stands for the Doppler compensated radar range measurement. $\mathbf{R} \in SO(3)$ is the rotation part of the transformation matrix in Equation (8). Thus, the uncertainty of the radar target in Cartesian coordinates can be derived as:

$$\hat{\Sigma} = \mathbf{J} \cdot \Sigma \cdot \mathbf{J}^{\text{T}}. \qquad (11)$$

Fig. 6 showcases the uncertainty estimation results of two radar scans. The gray areas around the measured radar targets

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

6   IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS
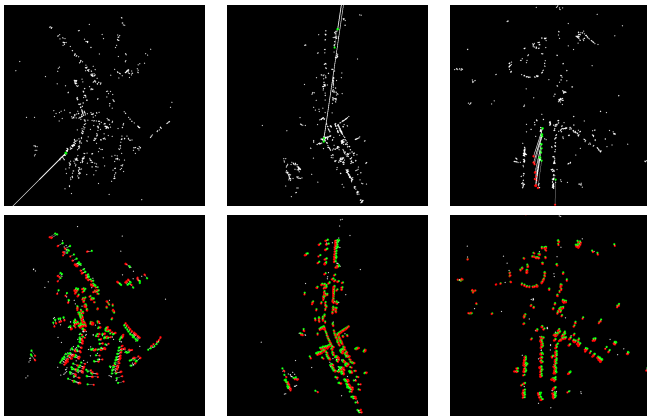
Fig. 7. The matching results between two radar sub-maps using RSD [26]: The green and red dots represent the points in the current and previous radar sub-maps respectively. The Doppler-effect-free radar sub-maps (bottom row) can achieve more robust matching compared with those with Doppler distortion (top row).
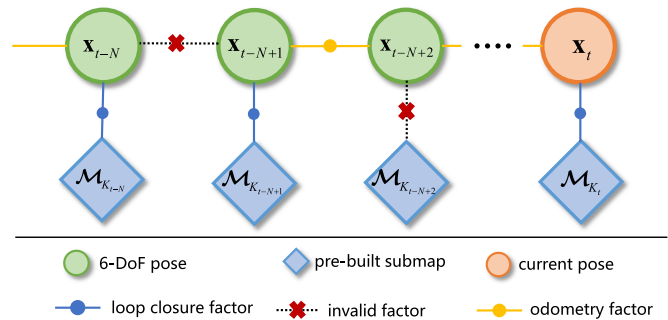

Fig. 8. Truncated pose graph structure: Each node represents the vehicle's state at a time or a sub-map observation at that time, and each edge represents a loop closure/odometry constraint. After estimating the relative pose of the current radar sub-map *w.r.t.* the pre-built map and the previous radar sub-map, a new state will add to the pose graph. To enhance the robustness, we filter out those localization results that significantly violate the pose prediction.

represent the possible uncertainty areas where the real targets may appear. As expected, those points further away from the sensor have larger uncertainties. Our system will punish these uncertain points by taking the Mahalanobis norm as part of the optimization objective.

2) Registration: Data association is the primary process for metric localization. In order to get good association results, we apply the state-of-the-art Radial Statistics Descriptor (RSD) [26] designed for radar point cloud to create a unique descriptor for each point in the sub-map. Then we take the *K-Nearest Neighbor* (KNN) algorithm to find coarse associations based on the RSD. Furthermore, RANSAC [27] excludes false matches from the initial associations. Fig. 7 shows that our Doppler-compensated radar sub-maps lead to more robust associations. The performance gain can be attributed to the Doppler compensation as the distortion destroys the geometric relationships between features, leading to inconsistent RSD to the same point in different sub-maps.

Given the query radar sub-map $\mathcal{M}_Q$, and a pre-built radar sub-map $\mathcal{M}_K$, we seek to find the best state $\mathbf{x} = \{\mathbf{q}, \mathbf{t}\}$, which can align the $\mathcal{M}_Q$ to $\mathcal{M}_K$, where $\mathbf{q}$ is quaternion, and $\mathbf{t} \in \mathbb{R}^3$ represent rotation and translation, respectively. Assuming we have N associated radar point pairs, for every point $\hat{\mathbf{p}}_i$ in $\mathcal{M}_Q$ and its corresponding point $\hat{\mathbf{q}}_i$ in $\mathcal{M}_K$, we adopt Generalized-ICP (GICP) [28] algorithm to get the optimal $\mathbf{x}^*$. The reason for using the GICP algorithm is that GICP extends the classical ICP algorithm by utilizing a distribution-to-distribution [29] matching approach, allowing it to account for the uncertainty of individual radar targets and weights each residual in Mahalanobis distance forms, *i.e.*:

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \frac{1}{2} \sum_{i=1}^{N} \mathbf{e}_i^{\mathrm{T}} \left(\hat{\Sigma}_{K_i} + \mathbf{R}\{\mathbf{q}\} \hat{\Sigma}_{Q_i} \mathbf{R}\{\mathbf{q}\}^{\mathrm{T}} \right)^{-1} \mathbf{e}_i,$$
$$\mathbf{e}_i = \hat{\mathbf{q}}_i - \left(\mathbf{R}\{\mathbf{q}\} \hat{\mathbf{p}}_i + \mathbf{t} \right). \tag{12}$$

$\hat{\Sigma}_{K_i}, \hat{\Sigma}_{Q_i}$ are uncertainties of $\hat{\mathbf{p}}_i, \hat{\mathbf{q}}_i$, derived from Sec. III-D.1. $\mathbf{R}\{\mathbf{q}\}$ denotes the rotation matrix corresponding to $\mathbf{q}$. We solve Equation. (12) using the Levenberg-Marquardt approach [30]. A Huber robust loss function [31] is also

integrated to make the loss less sensitive to outliers [32]. The transformation of the current body frame *w.r.t.* world frame (the coordinate when taking the first radar scan) $\mathbf{x}^w = \{\mathbf{q}^w, \mathbf{t}^w\}$ can be expressed as $\mathbf{x}^w = \mathbf{x}_r^w \boxplus \mathbf{x}^*$, with $\boxplus$ denoting the additive operation between two states, and $\mathbf{x}_r^w$ is the reference sensor's state in the world frame (*i.e.*, the state of the radar when constructing the pre-built map). For radar localization uncertainty modeling, we follow the method introduced in [33] to get the pose uncertainty $\mathbf{W}_{Q,K}$ of $\mathbf{x}$, *i.e.*,

$$\mathbf{W}_{Q,K} \approx \left(\frac{\partial^2 J}{\partial \mathbf{x}^2} \right)^{-1} \left(\frac{\partial^2 J}{\partial \mathbf{z} \partial \mathbf{x}} \right) cov(\mathbf{z}) \left(\frac{\partial^2 J}{\partial \mathbf{z} \partial \mathbf{x}} \right)^{\mathrm{T}} \left(\frac{\partial^2 J}{\partial \mathbf{x}^2} \right)^{-1},$$
$$J = \sum_{i=1}^{N} \| \hat{\mathbf{q}}_i - \left(\mathbf{R}\{\mathbf{q}\} \hat{\mathbf{p}}_i + \mathbf{t} \right) \|^2, \tag{13}$$

$\mathbf{z} = \left[\hat{\mathbf{p}}_1^{\mathrm{T}}, \hat{\mathbf{q}}_1^{\mathrm{T}}, \cdots, \hat{\mathbf{p}}_n^{\mathrm{T}}, \hat{\mathbf{q}}_n^{\mathrm{T}} \right]^{\mathrm{T}}$ is a vector consists of point correspondences between Q and K, and its covariance $cov(\mathbf{z})$ is

$$cov(\mathbf{z}) = \begin{bmatrix} \hat{\Sigma}_{K_1} & 0 & \cdots & 0 \\ 0 & \hat{\Sigma}_{Q_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{\Sigma}_{Q_n} \end{bmatrix}. \tag{14}$$

E. Truncated Pose Graph Optimization

The above registration process assumes that the input data is reliable to perform the metric localization. Unfortunately, due to the dynamic and complex scenes in practice (*e.g.*, in the presence of dynamic and occlusion objects), there exists less informative sensing data whose features are erroneous in space, causing wrong feature associations in radar maps and leading to large errors in localization. Such occasional but sudden performance degradation may result in fatal accidents in safety-critical applications such as autonomous driving. To improve the robustness of our system, we employ a truncated pose graph optimization scheme and devise a constraint validation criterion to detect and recover registration failures when there are less informative data.

1) Pose Graph Optimization: As illustrated in Fig. 8, we construct a local pose graph by incorporating loop closure and odometry constraints. Specifically, odometry constraints represent the relative transformation between two consecutive

radar frames, while loop closures represent the transformation between live radar scans and a pre-built radar map. All of the transformations are obtained according to Equation (12). We define the full state vector in the pose graph as:

$$\mathcal{X} = [\mathbf{x}_0, \mathbf{x}_1, \cdots \mathbf{x}_n],$$
$$\mathbf{x}_i = \left\{ \mathbf{q}_i^w, \mathbf{t}_i^w \right\}, \quad i = \{0, 1, 2, \cdots, n\}, \qquad (15)$$

where each state in the pose graph is expressed in the world frame. The relative pose estimation between states $\mathbf{x}_i$ and $\mathbf{x}_j$ is denoted as $\left\{ \hat{\mathbf{q}}_j^i, \hat{\mathbf{t}}_j^i \right\}$, which is derived from Sec. III-D.2. We denote the uncertainty of this relative pose as $\hat{\mathbf{W}}_{i,j}$. Given the states $\mathbf{x}_i$, $\mathbf{x}_j$ in the pose graph, and their relative pose estimates, we can add a factor into the pose graph by calculating the following residual:

$$\mathbf{r}_{i,j} \left(\mathbf{t}_i^w, \mathbf{q}_i^w, \mathbf{t}_j^w, \mathbf{q}_j^w \right) = \begin{bmatrix} \mathbf{R} \left\{ \mathbf{q}_i^w \right\}^{-1} \left(\mathbf{t}_j^w - \mathbf{t}_i^w \right) - \hat{\mathbf{t}}_j^i \\ 2 \left[\left(\mathbf{q}_i^w \right)^{-1} \otimes \mathbf{q}_j^w \otimes \left(\hat{\mathbf{q}}_j^i \right)^{-1} \right]_{xyz} \end{bmatrix}, \qquad (16)$$

where $[\cdot]_{xyz}$ indicates the vector part of a quaternion. We incorporate all the edges in the pose graph to construct the pose graph as shown in Fig. 8. Here we denote all the edges between consecutive states $\mathbf{x}_{i-1}, \mathbf{x}_i$ in $\mathcal{X}$ as $\mathcal{O}$, with $0 < i \leq n$. We put the edges between state $\mathbf{x}_i$ and its corresponding loop closure state $\mathbf{x}_k$ into a set denoted as $\mathcal{L}$. It is noted that loop closure state $\mathbf{x}_k$ is a constant vector and does not change in the subsequent calculations. All the states in the pose graph can thus be optimized by solving the following objective function:

$$\mathcal{X} = \arg \min_{\mathcal{X}} \sum_{i \in \mathcal{O}} \omega_{i-1,i} \left\| \mathbf{r}_{i-1,i} \right\|_{\hat{\mathbf{W}}_{i-1,i}}^2$$
$$+ \sum_{(k,i) \in \mathcal{L}} \omega_{k,i} \left\| \mathbf{r}_{k,i} \right\|_{\hat{\mathbf{W}}_{k,i}}^2, \qquad (17)$$

where $\omega \in \{0, 1\}$ is a binary state, denoting the validity of each factor in the pose graph.

2) Truncate Outliers: The key idea to improving the robustness of the system is to truncate invalid edges from the pose graph so that the pose graph will not optimize in the wrong direction. The basic idea to determine the validity of an edge is that the problematic radar scans will result in large pose estimation errors, denoted as $\hat{\mathbf{T}}_i^{i-1}$, in Sec. III-D.2. They are very likely to exhibit an obvious violation to the predictions of ego-velocity integration, denoted as $\bar{\mathbf{T}}_i^{i-1}$, in Sec. III-C.2. Below is the formulation for determining the validity of a measurement constraint. With ego-velocity integration estimation $\bar{\mathbf{T}}_i^{i-1}$ from Sec. III-C.2 and GICP estimated transformation $\hat{\mathbf{T}}_i^{i-1}$, $(i-1, i) \in \mathcal{O}$ and $\hat{\mathbf{T}}_i^k$, $(k, i) \in \mathcal{L}$ from Sec. III-D.2, we have

$$\Delta \xi_{i-1,i} = \log \left(\left(\hat{\mathbf{T}}_i^{i-1} \right)^{-1} \bar{\mathbf{T}}_i^{i-1} \right)^{\vee},$$
$$\Delta \xi_{k,i} = \log \left(\left(\hat{\mathbf{T}}_k^w \right)^{-1} \hat{\mathbf{T}}_{i-1}^w \bar{\mathbf{T}}_i^{i-1} \left(\hat{\mathbf{T}}_i^k \right)^{-1} \right)^{\vee}, \qquad (18)$$

where $\vee$ means converting skew-symmetric matrix to vector, and $\left\{ \Delta \xi_{i-1,i}, \Delta \xi_{k,i} \right\} \in \mathfrak{se}(3)$ represents relative pose deviations in Lie Group. We can further determine the validity of

both odometry and loop closure constraints by adopting the Chi-Squared test [34]:

$$D_e = (\Delta \xi_e)^{\mathrm{T}} \mathbf{W}_e^{-1} (\Delta \xi_e),$$
$$\omega_e = \begin{cases} 1 & \text{if} \quad D_e < T_e \\ 0 & \text{if} \quad D_e \geq T_e \end{cases}, \qquad (19)$$

where $\mathbf{W}_e$ represents the uncertainty of each registration result introduced in Sec. III-C.2. ω_e denotes the validity of each constraint. T_e is the acceptance threshold based on the Chi-Squared test with a significance level of 95%, which is a widely accepted threshold [35] because it enables effective filtering of localization anomalies while avoiding excessive false positives. The logic behind this criterion is that we expect that the pose estimation from a decent registration should not disagree with the pose prediction from ego-velocity integration too much. If some nodes lose edge connections after truncating, we use the integration of ego-velocities for compliments. Such a failure rejection mechanism makes `DC-Loc++` resilient to single registration failures.

IV. EVALUATION

Two datasets are used to evaluate the performance of `DC-Loc++`, including the real-world nuScenes dataset [11] and a synthetic dataset generated from the CARLA simulator [12]. In Sec. IV-A, we compare `DC-Loc++` with other mmWave radar-based metric localization methods. Then we perform an ablation study to reveal the functionality of each module and discuss the selection of the radar scans for the sub-map construction. In Sec. IV-B, we investigate the robustness of `DC-Loc++` under different velocity and noise settings on the synthetic CARLA dataset. In Sec. IV-C, we assess the real-time performance of the system by evaluating the efficiency of each module in `DC-Loc++`.

A. Performance Evaluation on nuScene

1) nuScenes Dataset: The nuScenes is an autonomous driving dataset collected by a 32-beam lidar, 5 cameras, 5 radars, and a GPS/IMU. The dataset has 242 km driving data in total at an average speed of 16 km/h. The radar used in this collection is the Continental ARS408-21 LRR automotive radar. This radar has a sampling rate of 13 Hz, 250 m sensing range, 0.39 m range resolution, 4.4° horizontal beamwidth, and 3.2° azimuth resolution. These settings are representative of today's automotive radars operating in the frequency band $76 \sim 77$ GHz. To obtain the sub-map pairs for evaluation, a place recognition process is typically required for loop closure detection. The optional place recognition algorithms include M2DP [36], AutoPlace [37], etc. However, since the place recognition algorithm is not our contribution, we resort to a simplified implementation, *i.e.*, the k-nearest neighbors algorithm, that uses the ground truth locations for the loop closure detection.

Besides, since the "Truncated Pose Graph Module" (c.f. Sec. III-E) needs multiple states and observations for pose graph construction, we only select the scenes with at least 3 sub-map pairs for testing.
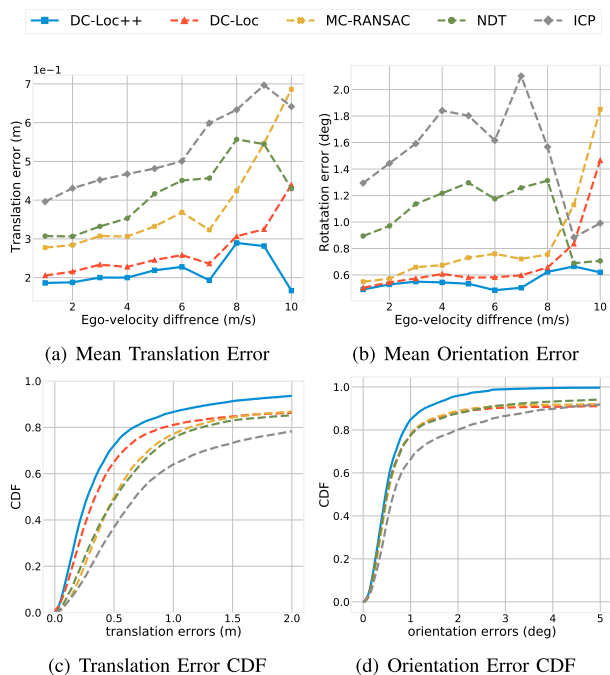
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

8

IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS

(a) Mean Translation Error (b) Mean Orientation Error

(c) Translation Error CDF (d) Orientation Error CDF

Fig. 9. Performance comparisons on the nuScenes dataset. Top row: mean translation and orientation error. Bottom row: cumulative distribution function (CDF) of localization results.

TABLE I
LOCALIZATION RESULTS ON THE NUSCENES DATASET

Methods	RTE (m)		RRE (deg)		SR (%)
	Mean	Median	Mean	Median	
ICP	0.529	0.398	1.760	0.683	84.8
NDT	0.415	0.351	1.249	0.506	89.6
MC-RANSAC	0.385	0.255	1.108	0.523	90.8
DC-Loc	0.269	0.181	0.964	0.489	90.2
DC-Loc++	**0.215**	**0.142**	**0.673**	**0.464**	**99.2**

In total, $11,995$ pairs of radar sub-maps are found valid for our purpose of metric localization.

2) Overall Performance on nuScenes: We compare our approach with 4 SOTA radar metric localization methods, including direct methods (convention ICP [21] and submap NDT [38]), and feature-based method (MC-RANSAC [8] and DC-Loc [10]). Note that the SOTA joint-Doppler-based NDT approach [39] is designed for odometry rather than metric localization. Thus we implement a grid-based NDT [38] for a fair comparison. Besides, since automotive radars do not have motion distortion, hence we only take the Doppler compensation module in MC-RANSAC for comparison. All methods use the radar sub-maps constructed by 10 consecutive radar scans. We set the same threshold of the translation error ($2.0\ m$) and rotation error ($5°$) to distinguish inliers and outliers for each method and only keep those inliers overall methods for performance analysis.

For ease of visualization, we present the results by grouping the velocity differences between two sensors and their corresponding translation and rotation errors on average. We compute the relative translation error (RTE[2]) and the relative rotation error (RRE[3]) for every method. In addition, we regard the pose as a successful estimation only if the RTE and RRE are below $2.0\ m$ and $5°$, and we use success rate (SR[4]) [2] to evaluate the stability of each method. Fig. 9 shows that the RTE of ICP and NDT tends to grow larger as the velocity difference between the two

---

[2]RTE $= \left(\frac{1}{m} \sum_{i=1}^{m} \|trans(E_i)\|^2 \right)^{\frac{1}{2}}$, where E_i is the relative transformation errors.

[3]RRE $= \left(\frac{1}{m} \sum_{i=1}^{m} \angle(rot(E_i)) \right)$, where $\angle(\cdot)$ represents for the degree of the rotation part in E_i.

[4]SR $= Results\_within\_thresholds/All\_results$.

sensors increases. We attribute their poor performance to the lack of addressing Doppler distortion and sparse radar outputs.

Notably, a sharp decrease is observed in RRE when the velocity differences are large ($\geq 8 m/s$). This is because ICP and NDT are sensitive to large transformations between the coordinate frames of radar scans. In other words, their metric localization accuracy substantially degenerates in the presence of large variations between live and pre-built radar maps. By investigating the transformations over different velocities in the nuScenes dataset, we find that the variations between radar sub-maps for high-velocity differences ($\geq 8\ m/s$) are substantially more minor, leading to better performance in the registration process.

MC-RANSAC achieves better performance compared with ICP and NDT, owing to its attempts to handle the Doppler effects. However, its accuracy is still inferior because it cannot completely correct the Doppler distortion without velocity observations. Thus, distorted radar sub-maps are used by the MC-RANSAC in metric localization. DC-Loc has the closest performance to our work. However, it lacks the capability to handle localization failures, making it vulnerable to less informative radar measurements. Moreover, when velocity differences surpass $8 m/s$, both MC-RANSAC and DC-Loc experience a deteriorated performance due to degenerated motions. Specifically, sub-map pairs with a large velocity difference ($\geq 8\ m/s$) are often collected at crossroads where a vehicle can be stationary at traffic lights in one instance and move at regular speeds in another instance. Since successive radar scans barely change when the vehicle is stationary, combining such radar scans will not provide additional information, making the sub-map remain spatially sparse. Recall that MC-RANSAC and DC-Loc are based on a descriptor-based registration method (RSD descriptor). The sub-map's sparsity highly degrades the descriptor's performance, leading to an accuracy decline in registration.

As shown in Table I, DC-Loc++ outperforms the compared methods in terms of all the metrics of interest. It reduces 44.1%, 20.8% average translation errors, and 25.3%, 17.7% average orientation errors compared with MC-RANSAC and DC-Loc, respectively. Furthermore, DC-Loc++ achieves a 9.0% improvement in the success rates and a lower variance compared with other works, as shown in Fig 10, indicating its ability to handle data association failures. The results demonstrate the superiority of DC-Loc++ in dealing with the Doppler effect and less informative radar measurements.

3) Ablation Study: We next investigate the individual contribution of each module in DC-Loc++. Table II shows the quantitative breakdowns.
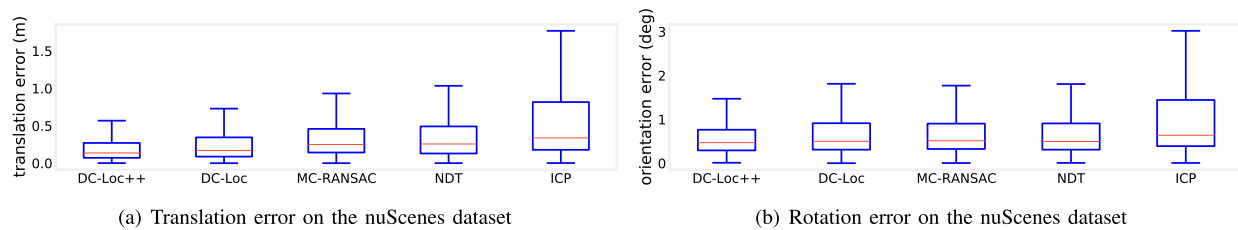
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

GAO et al.: ROBUST METRIC LOCALIZATION IN AUTONOMOUS DRIVING VIA DOPPLER COMPENSATION

9

(a) Translation error on the nuScenes dataset   (b) Rotation error on the nuScenes dataset

Fig. 10.   Quantitative analysis on nuScenes dataset. (a) and (b) are the RTE and RRE of each method. `DC-Loc++` has the lowest median error and variance compared with other methods.

TABLE II

ABLATION STUDY ON NUSCENES DATASET. ABBREVIATIONS: DOPPLER COMPENSATION (**D.**), UNCERTAINTY ESTIMATION (**U.**), TRUNCATED POSE GRAPH OPTIMIZATION (**T**)

Modules			RTE (m)	RRE (deg)	SR (%)
D.	**U.**	**T.**	Mean / Median	Mean / Median	
-	-	-	0.375 / 0.289	1.077 / 0.503	90.7
✓	-	-	0.297 / 0.191	1.033 / 0.495	90.7
-	✓	-	0.365 / 0.288	1.066 / 0.489	90.7
✓	✓	-	0.269 / 0.181	0.964 / 0.489	90.2
-	-	✓	0.309 / 0.265	0.683 / 0.468	98.8
✓	-	✓	0.220 / 0.146	0.678 / **0.463**	99.0
-	✓	✓	0.314 / 0.264	0.678 / 0.469	99.2
✓	✓	✓	**0.215 / 0.142**	**0.673** / 0.464	**99.2**

TABLE III

LOCALIZATION ERRORS ON NUSCENES DATASET WITH DIFFERENT NUMBERS OF CONCATENATED RADAR SCANS

No.	RTE (m)	RRE (deg)	SR (%)	Time (ms)
	Mean / Median	Mean / Median		
1	0.463 / 0.314	0.990 / 0.543	96.2	**10.4**
3	0.314 / 0.185	0.772 / 0.480	98.9	42.3
5	0.261 / 0.161	0.691 / 0.466	98.9	80.6
8	0.225 / 0.151	0.675 / **0.454**	99.2	156.5
10	**0.215 / 0.142**	**0.673** / 0.464	99.2	202.6
13	0.240 / 0.151	0.696 / 0.490	99.3	374.9
15	0.247 / 0.154	0.712 / 0.496	**99.4**	411.7

- **Doppler Compensation** (c.f. Sec. III-C): We can see that the proposed Doppler compensation alone can significantly improve the average translation and rotation estimation performance. The mean and median translation errors have been reduced by 20.8% and 33.9%. Such performance is slightly better than SOTA Doppler compensation solutions [10]. The barrier that prevents further accuracy improvement is the unreliability of mmWave sensing.
- **Uncertainty Estimation** (c.f. Sec. III-D): To better handle the unreliability of mmWave radar, we estimate the measurement uncertainty and take it as a confidence weight to the point cloud. Intuitively, our solution gives more weight to those points with small uncertainties. This treatment leads to 9.4% and 6.7% error reductions in translation and rotation estimation. However, tracking the measurement uncertainty and the Doppler compensation above cannot improve the success rate of metric localization. The failure cases (about 9.8% of trials) are due to the incorrect data associations from less informative radar scans.
- **Truncated Pose Graph Optimization:** (c.f. Sec. III-E): Clearly, the truncated pose graph optimization alone can raise the success rate to 98.8%, making our system insensitive to less informative outliers. It is attributed to the multi-frame constraints and our edge validity determination schemes, which can mitigate the adverse impacts of localization anomalies (failure edges) in the pose graph and use other valid edges to ensure localization reliability.

In a nutshell, the proposed modules have complementary functionalities. Specifically, the Doppler compensation module significantly improves localization accuracy by correcting the Doppler distortions. The uncertainty estimation module can further improve the accuracy of the system by accounting for the uncertainty of each radar target. The truncated pose graph optimization module enhances the robustness by correcting metric localization anomalies. So there is no doubt that the best performance is achieved by using three modules in tandem.

4) Choice for the Number of Concatenated Radar Scans: The number of radar scans used for sub-map construction affects the accuracy and efficiency of the system. In order to determine the optimal number of radar scans for sub-map construction, we test DC-Loc++ with different numbers of radar scans. As shown in Table III, the optimal performance can be achieved when concatenating 10 successive radar scans. When we use fewer than 10 radar scans, the radar sub-maps cannot provide adequate radar targets required for data association, degrading the performance. In contrast, concatenating excessive radar scans (≥ 10) will introduce higher cumulative noise levels and increase the computational cost, leading to a notable reduction in localization accuracy and efficiency. Thus, using 10 successive radar scans for sub-map construction is optimal for both the accuracy and efficiency of the system.

B. Robust Analysis on CARLA

Most of the nuScenes dataset is collected at low ego-velocities, and the parameters of the onboard radar cannot be changed either. Therefore, it can hardly comprehensively demonstrate the robustness of our method in various settings. To fully understand the robustness under various ego-velocities and noise models, we synthesize more data by tuning radar parameters through the CARLA simulator.

1) Data Collection: Specifically, a car with a forward-facing radar is rendered in our experiment to collect data for localization. The maximum detection range and the radar's field of view (FOV) are set to be 100 m and 150°, respectively. To examine different noise models of radar measurements, we generate four datasets with the noise following Gaussian
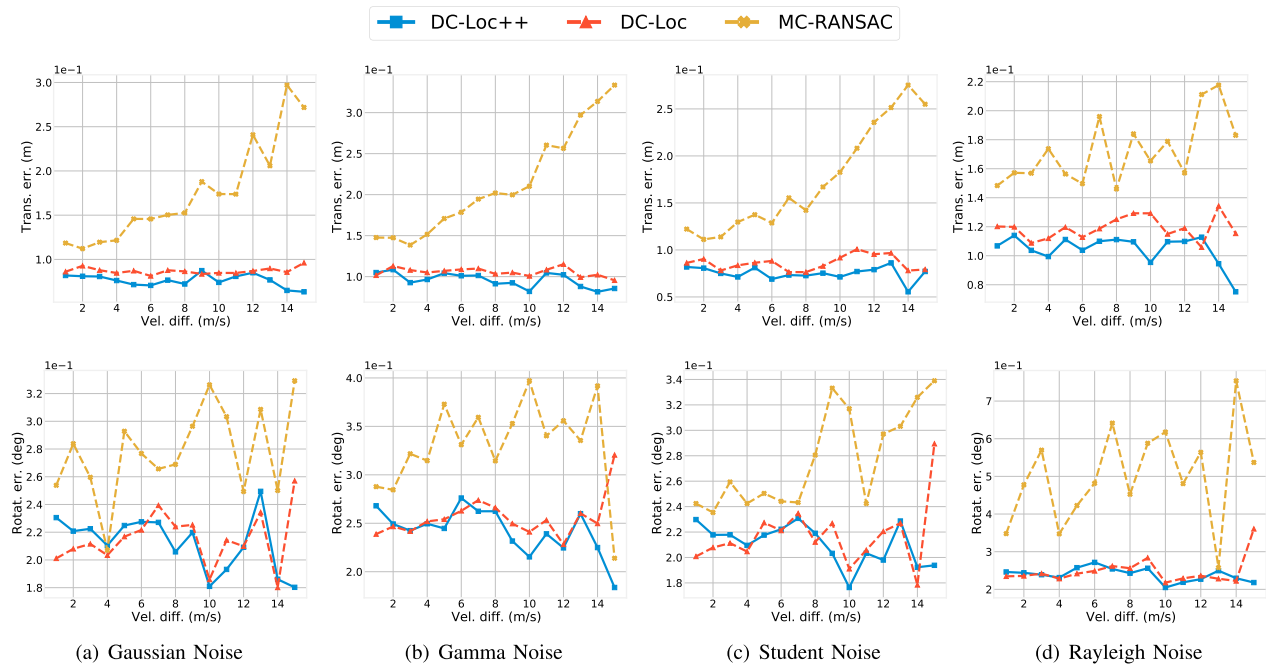
Fig. 11. The median translation and rotation errors on the CARLA dataset with different noise models. From left to right: Gaussian noise, Gamma noise, Student noise, and Rayleigh noise.

TABLE IV
NOISE MODELS IN THE SYNTHETIC CARLA DATASET

Noise	Range	Angle	Velocity
Gaussian	$\mathcal{N}(0, 0.25^2)$	$\mathcal{N}(0, 0.5^{\circ 2})$	$\mathcal{N}(0, 0.1^2)$
Gamma	$\Gamma(0, 0.25)$	$\Gamma(0, 0.025^\circ)$	$\Gamma(0, 0.1)$
Student	$0.25 \times t(100)$	$0.5^\circ \times t(100)$	$0.1 \times t(100)$
Rayleigh	$Rayleigh(0.5)$	$Rayleigh(1)$	$Rayleigh(0.1)$

distribution, Gamma distribution, Student's t-distribution, and Rayleigh distribution, respectively. In particular, the noise to the range, angle, and velocity measurements are listed in Table IV. We simulate the Doppler effect with $\frac{f_c}{K} = 0.04$ (c.f. Equation (5)), following the technical parameters of ARS408-21. We conducted two rounds of radar data collection to create the sub-map pairs for metric localization. The average speed of the former is 72 km/h, and the latter is 40 km/h, featuring a larger ego-velocity difference than that of nuScenes (around trifold larger). After applying the k-nearest neighbors algorithm for the loop closure detection, 3,130 pairs of loop closures are eventually created for evaluation.

2) Robust Analysis: We compare our approach with MC-RANSAC [8] and DC-Loc [10] on the aforementioned synthetic dataset and plot the median translation and rotation errors in Fig. 11. As we can see, the translation and rotation errors of MC-RANSAC (the baseline second to the best on nuScenes) have a pronounced tendency to increase as the differences between ego-velocities become larger. In contrast, DC-Loc++ is robust to the Doppler effect and achieves a consistent performance under variant ego-velocities. Furthermore, our method has good stability under the effect of different types of noise. Specifically, DC-Loc++ can outperform the MC-RANSAC by reducing 51.9%, 53.4%,

52.7%, 37.5% translation errors and 29.8%, 31.0%, 23.9%, 53.2% rotation errors with Gaussian noise, Gamma noise, Student's noise, and Rayleigh noise, respectively. In addition, compared with our previous work [10], we can achieve 15.1%, 11.8%, 12.3%, 13.4% translation improvement, and comparable orientation performance with noise described above. A localization trajectory is presented in Fig. 12. We can observe that DC-Loc [10] and MC-RANSAC [8] produce a zigzag trajectory and even fail in localization occasionally. In contrast, DC-Loc++ bounds the localization error within $0.48m$, which means our method can achieve a more robust and accurate localization performance. This can be attributed to DC-Loc++'s capability of utilizing multi-frame constraints and correcting localization failures.

C. Efficiency Evaluation

To assess the time cost of our proposed radar localization system, we measure the computation efficiency on both nuScenes dataset and the synthetic Carla dataset. More than 15,000 radar corresponding pairs are fed into our system. All the experiments are conducted with an Intel i5-10300H CPU operating at 2.5GHz, and the average time cost of each module is presented in Table V. The total time cost for every estimation is $260.1ms$, and the whole system runs at 3.85 Hz on average. Since the ARS408-21 FMCW radar operates at 13 Hz, our implementation is currently not real-time capable. However, we can use inter-frame ego-velocities to predict the poses, leading to an equally real-time performance at approximately 17.7 Hz. Although such an operation may compromise localization accuracy to some extent, the "Truncated Pose Graph" module (c.f. Sec. III-E) can refine the results at the arrival of the next radar sub-map.
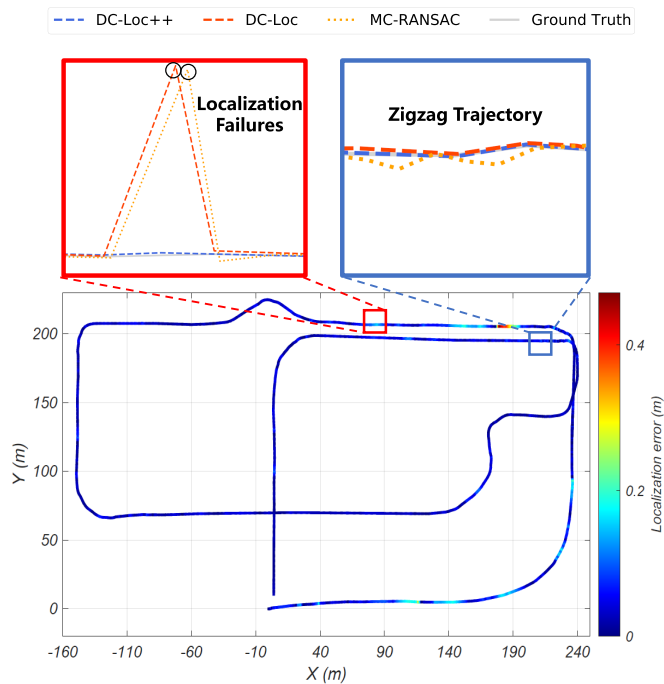
Fig. 12. Qualitative results of `DC-Loc++` on synthetic Carla dataset, including zoomed views at square areas. As the metric localization errors are relatively smaller than other localization tasks (*e.g.*, odometry), we omit the full trajectory of the compared methods and only plot part of them in the zoomed views for better visualization. We can observe that `DC-Loc++` achieves a more robust performance than other methods by detecting and recovering localization failures (top left). Moreover, `DC-Loc++` provides a more smooth and more accurate trajectory by using Doppler compensation and uncertainty estimation algorithms (top right).

TABLE V
TIME CONSUMPTION ON EACH MODULE

Modules	D.C.	U.R.	T.P.	Total
Time (ms)	56.8	202.6	0.7	260.1

D.C. - Doppler Compensation
U.R. - Uncertainty-aware Registration
T.P. - Truncated Pose Graph Optimization

In `DC-Loc++`, the "Uncertainty-aware Registration" module (c.f. Sec. III-D) is the most time-consuming module where the bottleneck is the iterative process of finding corresponding inlier pairs between radar scans. We believe a more efficient implementation can be achieved by adopting some specialized outlier rejection algorithms *e.g.*, [40]. We leave the implementation of such algorithms as future work and will not discuss them further in this paper. It is worth noting that the truncated pose graph module can significantly enhance the overall estimation accuracy and robustness with only marginal time ($0.7ms$) in computation, which is very cost-efficient for the whole system.

V. CONCLUSION

In this paper, an automotive radar-based framework `DC-Loc++` is proposed to tackle the fundamental problems of the Doppler effect and noisy radar measurements in radar metric localization. We first propose an explicit Doppler compensation method in the radar mapping process to combat Doppler distortion in radar scans. Then we model the uncertainty of each radar target, taking into account both range and Doppler velocity variance to mitigate the impact of radar artifacts. Finally, a truncated pose graph optimization scheme is designed to detect and recover individual localization failures to make our method more reliable. Extensive experiments using the real-world data from nuScenes and the synthetic data from CARLA demonstrate the effectiveness of our method in a variety of scenarios by reducing more than 20.0% and 30.2% errors in terms of translation and rotation.

In our future work, we will improve the computational efficiency and localization accuracy of the proposed method. In addition, we will integrate our method into an automotive radar-based SLAM system.

REFERENCES

[1] P. Agarwal, W. Burgard, and L. Spinello, "Metric localization using Google Street View," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2015, pp. 3111–3118, doi: 10.1109/IROS.2015.7353807.

[2] L. Luo, S.-Y. Cao, B. Han, H.-L. Shen, and J. Li, "BVMatch: LiDAR-based place recognition using bird's-eye view images," *IEEE Robot. Autom. Lett.*, vol. 6, no. 3, pp. 6076–6083, Jul. 2021, doi: 10.1109/LRA.2021.3091386.

[3] D. Barnes and I. Posner, "Under the radar: Learning to predict robust keypoints for odometry estimation and metric localisation in radar," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 9484–9490, doi: 10.1109/ICRA40945.2020.9196835.

[4] T. Y. Tang, D. De Martini, S. Wu, and P. Newman, "Self-supervised localisation between range sensors and overhead imagery," 2020, *arXiv:2006.02108*.

[5] W. Wang, P. P. B. de Gusmão, B. Yang, A. Markham, and N. Trigoni, "RadarLoc: Learning to relocalize in FMCW radar," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2021, pp. 5809–5815, doi: 10.1109/ICRA48506.2021.9560858.

[6] D. De Martini, M. Gadd, and P. Newman, "KRadar++: Coarse-to-fine FMCW scanning radar localisation," *Sensors*, vol. 20, no. 21, p. 6002, Oct. 2020, doi: 10.3390/s20216002.

[7] M. I. Skolnik, *Radar Handbook*. New York, NY, USA: McGraw-Hill, 2008.

[8] K. Burnett, A. P. Schoellig, and T. D. Barfoot, "Do we need to compensate for motion distortion and Doppler effects in spinning radar navigation?" *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 771–778, Apr. 2021, doi: 10.1109/LRA.2021.3052439.

[9] K. Burnett, Y. Wu, D. J. Yoon, A. P. Schoellig, and T. D. Barfoot, "Are we ready for radar to replace LiDAR in all-weather mapping and localization?" *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 10328–10335, Oct. 2022, doi: 10.1109/LRA.2022.3192885.

[10] P. Gao, S. Zhang, W. Wang, and C. X. Lu, "DC-Loc: Accurate automotive radar based metric localization with explicit Doppler compensation," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2022, pp. 4128–4134, doi: 10.1109/ICRA46639.2022.9811561.

[11] H. Caesar et al., "nuScenes: A multimodal dataset for autonomous driving," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11618–11628, doi: 10.1109/CVPR42600.2020.01164.

[12] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in *Proc. 1st Annu. Conf. Robot Learn.*, in Proceedings of Machine Learning Research, vol. 78, S. Levine, V. Vanhoucke, and K. Goldberg, Eds., Nov. 2017, pp. 1–16.

[13] X. Ding et al., "Persistent stereo visual localization on cross-modal invariant map," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 11, pp. 4646–4658, Nov. 2020, doi: 10.1109/TITS.2019.2942760.

[14] Y. Zhang, H. Zhang, G. Wang, J. Yang, and J.-N. Hwang, "Bundle adjustment for monocular visual odometry based on detections of traffic signs," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 151–162, Jan. 2020, doi: 10.1109/TVT.2019.2954876.

[15] J. Yu, Z. Xiang, and J. Su, "Hierarchical multi-level information fusion for robust and consistent visual SLAM," *IEEE Trans. Veh. Technol.*, vol. 71, no. 1, pp. 250–259, Jan. 2022, doi: 10.1109/TVT.2021.3127554.

[16] E. Javanmardi, M. Javanmardi, Y. Gu, and S. Kamijo, "Pre-estimating self-localization error of NDT-based map-matching from map only," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 12, pp. 7652–7666, Dec. 2021, doi: 10.1109/TITS.2020.3006854.

[17] T. Shan and B. Englot, "LeGO-LOAM: Lightweight and ground-optimized LiDAR odometry and mapping on variable terrain," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 4758–4765, doi: 10.1109/IROS.2018.8594299.

[18] W. Lu, Y. Zhou, G. Wan, S. Hou, and S. Song, "L^3-Net: Towards learning based LiDAR localization for autonomous driving," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 6382–6391, doi: 10.1109/CVPR.2019.00655.

[19] W. Lu, G. Wan, Y. Zhou, X. Fu, P. Yuan, and S. Song, "DeepVCP: An end-to-end deep neural network for point cloud registration," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 12–21, doi: 10.1109/ICCV.2019.00010.

[20] E. Ward and J. Folkesson, "Vehicle localization with low cost radar sensors," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2016, pp. 864–870, doi: 10.1109/IVS.2016.7535489.

[21] P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," *Proc. SPIE*, vol. 1611, pp. 586–606, Jan. 1992, doi: 10.1117/12.57955.

[22] Z. Hong, Y. Petillot, and S. Wang, "RadarSLAM: Radar based large-scale SLAM in all weathers," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2020, pp. 5164–5170, doi: 10.1109/IROS45743.2020.9341287.

[23] H. Yin, R. Chen, Y. Wang, and R. Xiong, "RaLL: End-to-end radar localization on LiDAR map using differentiable measurement model," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 6737–6750, Jul. 2022, doi: 10.1109/TITS.2021.3061165.

[24] S. Rao, "Introduction to mmWave sensing: FMCW radars," Texas Instrum., Austin, TX, USA, Tech. Rep., 2017. [Online]. Available: https://www.ti.com/content/dam/videos/external-videos/2/3816841626001/5415528961001.mp4/subassets/mmwaveSensing-FMCW-offline viewing_0.pdf

[25] D. Kellner, M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer, "Instantaneous ego-motion estimation using multiple Doppler radars," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2014, pp. 1592–1597, doi: 10.1109/ICRA.2014.6907064.

[26] S. H. Cen and P. Newman, "Radar-only ego-motion estimation in difficult settings via graph matching," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 298–304, doi: 10.1109/ICRA.2019.8793990.

[27] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981, doi: 10.1145/358669.358692.

[28] A. Segal, D. Hähnel, and S. Thrun, "Generalized-ICP," in *Robotics: Science and Systems*, J. Trinkle, Y. Matsuoka, and J. A. Castellanos, Eds. Cambridge, MA, USA: MIT Press, 2009, doi: 10.15607/RSS.2009.V.021.

[29] K. Koide, M. Yokozuka, S. Oishi, and A. Banno, "Voxelized GICP for fast and accurate 3D point cloud registration," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2021, pp. 11054–11059, doi: 10.1109/ICRA48506.2021.9560835.

[30] J. J. Moré, "The Levenberg–Marquardt algorithm: Implementation and theory," in *Numerical Analysis*. Cham, Switzerland: Springer, 2006, pp. 105–116, doi: 10.1007/BFb0067700.

[31] P. J. Huber, "Robust estimation of a location parameter," in *Breakthroughs in Statistics*. Cham, Switzerland: Springer, 1992, pp. 492–518, doi: 10.1214/aoms/1177703732.

[32] J. T. Barron, "A general and adaptive robust loss function," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4326–4334, doi: 10.1109/CVPR.2019.00446.

[33] S. M. Prakhya, L. Bingbing, Y. Rui, and W. Lin, "A closed-form estimate of 3D ICP covariance," in *Proc. 14th IAPR Int. Conf. Mach. Vis. Appl. (MVA)*, May 2015, pp. 526–529, doi: 10.1109/MVA.2015.7153246.

[34] M. L. McHugh, "The chi-square test of independence," *Biochemia Medica*, vol. 23, no. 2, pp. 143–149, 2013, doi: 10.11613/BM.2013.018.

[35] R. A. Fisher, *Statistical Methods for Research Workers*. New York, NY, USA: Springer, 1954.

[36] L. He, X. Wang, and H. Zhang, "M2DP: A novel 3D point cloud descriptor and its application in loop closure detection," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 231–237, doi: 10.1109/IROS.2016.7759060.

[37] K. Cai, B. Wang, and C. X. Lu, "AutoPlace: Robust place recognition with single-chip automotive radar," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2022, pp. 2222–2228, doi: 10.1109/ICRA46639.2022.9811869.

[38] P. Biber and W. Strasser, "The normal distributions transform: A new approach to laser scan matching," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Mar. 2003, pp. 2743–2748, doi: 10.1109/IROS.2003.1249285.

[39] P.-C. Kung, C.-C. Wang, and W.-C. Lin, "A normal distribution transform-based radar odometry designed for scanning and automotive radars," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2021, pp. 14417–14423, doi: 10.1109/ICRA48506.2021.9561413.

[40] K. Cai, C. X. Lu, and X. Huang, "Uncertainty estimation for 3D dense prediction via cross-point embeddings," 2022, *arXiv:2209.14602*.

Pengen Gao received the B.S. degree from the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan, in 2020, where he is currently pursuing the M.S. degree. His research interests include SLAM and radar-based localization.

Shengkai Zhang (Member, IEEE) received the Ph.D. degree from the School of Electronic Information and Communications, Huazhong University of Science and Technology, China, in 2021. He is currently an Associate Professor with the Wuhan University of Technology, China. His current research interests include wireless sensing, localization, mobile computing, multi-sensor fusion, and robot control and planning.

Wei Wang (Senior Member, IEEE) received the Ph.D. degree from the Department of Computer Science and Engineering, The Hong Kong University of Science and Technology. He is currently a Full Professor with the School of Electronic Information and Communications, Huazhong University of Science and Technology. He has published two books and 90 refereed papers in international leading journals and primer conferences. He is the inventor of three U.S. and 20 Chinese patents. His research interests include PHY/MAC design and mobile computing in wireless systems. He has won the Best Paper Award in IEEE ICC 2019. He was selected for the Young Elite Scientist Sponsorship Program, China Association for Science and Technology; and the Hundred Talents Program, Hubei, China. He served on TPC for INFOCOM, GBLOBECOM, and ICC. He served as an Editor for IEEE ACCESS, *IJCS*, and *China Communications*; and the Guest Editor for *Wireless Communications and Mobile Computing* and the IEEE COMSOC MMTC COMMUNICATIONS.

Chris Xiaoxuan Lu received the M.Eng. degree from Nanyang Technology University, Singapore, and the Ph.D. degree from the University of Oxford. He is an Assistant Professor with the School of Informatics, The University of Edinburgh. His research interest lies in mobile robotics, cyber physical systems, and machine perception.