A Compressed Sensing and Clustering Method for Robotic Localization using a Single FMCW Radar

Zelin Hao, Siyuan Zhao, Zhi Zheng, Zhiyong Huang, Yongxin Guo*
Department of Electrical and Computer Engineering National University of Singapore
*eleguoyx@nus.edu.sg

***Abstract*—This paper presents a robust and precise ego-motion estimation system based on millimeter wave (mmWave) frequency-modulated continuous wave (FMCW) radar. Compressed sensing is introduced to greatly enhance radar point cloud resolution. The designed landmark association algorithm leverages cluster-based features and Doppler velocity to ameliorate registration in challenging scenarios, thereby fortifying the robustness of the system. Experimental results demonstrate that the proposed algorithm achieves high-quality point cloud generation and precise ego-motion estimation using a single commercial radar.**

***Index Terms*—compressed sensing, clustering, ego-motion estimation, FMCW radar, radar point cloud**

I. Introduction

With the increasing prevalence of autonomous vehicles and mobile robots, precise localization technologies have emerged as a crucial research area. It plays a vital role in ensuring the safe and efficient execution of tasks for robots, particularly in challenging environments such as extreme weather, disaster relief operations, and geological exploration. However, ego-motion estimation, a key technique in robotic localization, relying on optical sensors often encounters inherent limitations due to variations in light intensity and air visibility [1]. Fortunately, advancements in integrated circuit technologies have paved the way for a new possible solution: low-cost, compact, all-weather operating conditions, and high-performance mmWave imaging radars. Leveraging these advancements, mmWave imaging radars offer an alternative with robust performance in a wide range of scenarios.

However, the utilization of mmWave radar for precise ego-motion estimation still encounters various challenges. The raw radar data demonstrates a diminished signal-to-noise ratio (SNR), limited resolution, and vulnerability to multipath effects. Nevertheless, after undergoing cleaning procedures, each scan provides only limited information. In comparison to lidar and camera data, the filtered point cloud consists of only hundreds of points, showcasing a notable sparsity. Recent endeavors have focused on addressing these challenges in radar ego-motion estimation. Some approaches have explored the use of deep learning techniques to improve data quality and estimation accuracy [2]. Nevertheless, these deep learning-based methods suffer from poor transferability in unfamiliar scenes and require significant training costs. Alternatively, [3] integrated additional sensors for higher performance while coming at the expense of increased hardware costs and heightened system complexity.

To suppress the defects of existing methods, this paper proposes a high-precision and robust radar point cloud generation and registration algorithm. In this paper, the utilization of compressed sensing improves radar point cloud resolution without offline learning; the landmark association algorithm extracts valid landmark matches by applying cluster-based features and Doppler information, eliminating random noise and false detection. Experimental results demonstrated that the proposed method enables precise pose estimation and motion trajectory, thereby establishing a solid foundation for robotic localization

II. Approach

A. High-Resolution Radar Point Cloud Generation

The lack of angular resolution caused by the limitation of antenna aperture is one of the biggest obstacles to the realization of high-precision robot positioning using mmWave radar. The random movement also makes it challenging to implement Synthetic-aperture radar (SAR) in the system. Fortunately, for linear array multiple-input and multiple-output (MIMO) radar, compressed sensing (CS) is leveraged to achieve angular super-resolution radar images [4]. For a MIMO array consisting of N_T transmit antennas and N_R receive antennas, the $N_T \times N_R$ virtual array elements can be realized through Time Division Multiplexing (TDM) MIMO. The steering vector, which represents the phase delays on each transmit-receive pair, could be represented as

$$\mathbf{a}(\theta) = \begin{bmatrix} e^{-j\frac{2\pi}{\lambda}d_T\sin(\theta)} \times e^{-j\frac{2\pi}{\lambda}d_R\sin(\theta)} \\ \vdots \\ e^{-j\frac{2\pi}{\lambda}d_T\sin(\theta)} \times e^{-j\frac{2\pi}{\lambda}N_R d_R\sin(\theta)} \\ \vdots \\ e^{-j\frac{2\pi}{\lambda}N_T d_T\sin(\theta)} \times e^{-j\frac{2\pi}{\lambda}N_R d_R\sin(\theta)} \end{bmatrix} \quad (1)$$

where θ represents the azimuth of the incident wave. When there are different azimuth targets, the signal received by the virtual array can be represented as

$$\mathbf{s} = \sum_{k=1}^{K} \chi_k \mathbf{a}(\theta_k) + \mathbf{w} \quad (2)$$

where $\mathbf{s} \in \mathbb{R}^{N_T \times N_R}$ represents the radar echo on all virtual channel, and χ_k represents the reflection intensity of the kth target and $\mathbf{w} \in \mathbb{R}^{N_T \times N_R}$ is additive white Gaussian noise. The received signal is composed of signals encompassing different frequencies with noise, and each frequency component corresponds to one azimuth angle. The direction of arrival (DOA) is usually computed after the distance Fast Fourier Transform (FFT) estimation and the range-velocity 2D-FFT estimation in the FMCW radar system. The conventional DOA estimation is obtained through matched filtering method by performing 3D-FFT on the range-Doppler matrix along all channels. However, it often results in high sidelobes, which directly leads to low angular resolution and noisy data, due to the limitation of antenna aperture and random noise. Fortunately, the signal can be assumed to be sparse, given the rarity of targets with the same range and velocity. Therefore, CS could be applied for super-resolution radar images generation, where azimuth estimation can be converted into a special type of sparse linear inverse problem:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{3}$$

where $\mathbf{y} \in \mathbb{R}^D$ is the received signal which is padded to a power of 2 for efficient radix-2 butterflies operation, $D = ceiling(log2(N_T \times N_R))$. The sensing matrix $\mathbf{A} \in \mathbb{R}^{D\times D}$ is determined by the imaging system parameters and geometry. In our system, $\mathbf{A}$ stands for inverse Fourier transform matrix. $\mathbf{x} \in \mathbb{R}^D$ is the azimuth of the target to be recovered and $\mathbf{n} \in \mathbb{R}^D$ models the noises and disturbances. The goal is to recover the reflection coefficients and azimuth of the scattered points under the premise of known radar echoes and measurement matrices. The problem is an optimization problem as

$$\hat{\mathbf{x}} = \arg\min_x \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \epsilon\|\mathbf{x}\|_1, \tag{4}$$

where ϵ is a tunable parameter that controls the trade-off between sparsity and reconstruction error constraint. Due to its remarkable convergence capability, Approximate Message Passing (AMP) is utilized to address this optimization problem, which iterates the steps (for $t = 0, 1, 2, ..., x_0 = A^H y$, and $v_0 = y$):

$$\begin{cases} \mathbf{v}_t = \mathbf{y} - \mathbf{A}\mathbf{x}_t + b_t \mathbf{v}_{t-1} \\ \mathbf{x}_{t+1} = \eta_{\mathrm{st}}\left(\mathbf{x}_t + \mathbf{A}^H \mathbf{v}_t, \tau_t\right) \\ b_t = \frac{1}{M}\left\|\mathbf{x}_t\right\|_0 \\ \tau_t = \frac{\kappa}{\sqrt{M}}\left\|\mathbf{v}_t\right\|_2 \\ \left[\eta_{\mathrm{st}}\left(\mathbf{r}, \tau_t\right)\right]_i = \operatorname{sgn}\left(r_i\right)\max\left\{\left|r_i\right| - \tau_t, 0\right\} \end{cases} \tag{5}$$

where v_t means the residual measurement error and κ is a tuned parameter [5]. Super-resolution radar images can be obtained by performing the AMP algorithm on all the range bins where targets exist. The radar point cloud is then filtered out when the intensity of the image pixel is above the dynamic threshold.

B. Clustering Landmark Association

Our proposed clustering landmark association algorithm is designed to optimize registration for a robust system. High-precision matching requires point clouds with high density and fewer outliers. However, the multipath effects introduce numerous outliers and decrease the instability of the same target in consecutive frames. Considering this issue, a clustering landmark association algorithm has been proposed for scan matching. Firstly, the density-based spatial clustering of application with noise (DBSCAN) algorithm has been applied to cluster the point cloud P and filters out clusters with low average intensity or few points. The remaining clusters, denotes as landmarks $L^{1\ldots M}$, are used to represent targets.

Algorithm 1: Landmark Feature Generation Method

Function CreateFeature(L,g,r)
- **for** $L^m \in L$ **do**
  - $key_{L^m} \leftarrow [max(p_I^m); arithmMean(p_I^m); harmMean(p_I^m)];$
  - $LocalSection \leftarrow sort(p_I^m);$
  - $shift(LocalSection);$
- **for** $L^i \in L$ **do**
  - **for** $L^j \in L$ **do**
    - $index := \left\lceil \frac{\left\|(p^i_{max_x}, p^i_{max_y}) - (p^j_{max_x}, p^j_{max_y})\right\|_2}{r/g} \right\rceil + 1$;
    - $GlobalSection(index)++;$

Algorithm.1 summarizes the generation of the landmark feature descriptor $d_{L^m} : (GlobalSection, LocalSection)$ and representative point key_{L^m} for $L^m : \{p_1^m, p_2^m, ..., p_N^m\} \in L$. Inspired by [6], the $GlobalSection$ captures the distance information and distribution relationship among landmarks. Subsequently, Pearson correlation coefficient ρ between two $GlobalSections$ Gs_1 and Gs_2 from two landmark sets L_1 and L_2 is calculated by

$$\rho\left(Gs_1, Gs_2\right) = \frac{\sum\left(Gs_1 - \overline{Gs_1}\right)\left(Gs_2 - \overline{Gs_2}\right)}{\sqrt{\sum\left(Gs_1 - \overline{Gs_1}\right)^2 \sum\left(Gs_2 - \overline{Gs_2}\right)^2}} \tag{6}$$

Given the Gaussian-like characteristics of intensity distribution, the Wasserstein distance is employed to compute the correlation between the two $LocalSections$ Ls_1 and Ls_2 by

$$W\left(Ls_1, Ls_2\right) = inf_{\gamma \in S(Ls_1, Ls_2)} E\left(x, y\right)\gamma\left[\|x - y\|\right] \tag{7}$$

where S represents all possible joint distributions between the given distributions and each joint distribution γ is utilized to characterize the cost of transformation between these distributions. The landmarks exhibiting a substantial correlation are then incorporated into $candidates_1$ and $candidates_2$ based on ρ and W respectively. Each landmark L^m possesses its corresponding $candidates_1$ and $candidates_2$. Duplicated landmarks with the closest Euclidean distance between their respective candidates are added to the Matches M.

The radar's translation velocity can be used to further remove mismatches in polar coordinates. It can be computed by

$$\begin{bmatrix} \cos\theta_1 & \sin\theta_1 \\ \cos\theta_2 & \sin\theta_2 \\ \vdots & \vdots \\ \cos\theta_N & \sin\theta_N \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} -\dot{r}_1 \\ -\dot{r}_2 \\ \vdots \\ -\dot{r}_N \end{bmatrix} \tag{8}$$