Robust Spatial Perception with 4D Radar for Mobile Autonomy

Fangqiang Ding, School of Informatics, University of Edinburgh

I. Introduction

A critical pillar for achieving mobile autonomy is robust spatial perception—the ability to accurately understand the ambient environment and precisely localize ego-agents under diverse and complex scenarios [1, 22]. Traditional spatial perception approaches heavily rely on optical sensors, *i.e.,* depth/RGB camera and LiDAR [17, 19], which lack robustness against adverse weather (*e.g.,* fog, rain, snow) and bad lighting conditions (*e.g.,* darkness, sun glare) [3, 23]. This vulnerability poses significant risks to the safety of autonomous agents, limiting their widespread, long-term deployment in the wild.

Inspired by biological sensing mechanisms which leverage non-visual cues for hunting and navigation [20, 12], my research explores 4D single-chip mmWave radar [16, 8] (*i.e.,* 3D position + Doppler velocity, hence '4D'), as a complementary sensor modality for spatial perception systems. As an emerging sensor technology, 4D radar is reputable for its robustness against adverse weathers, cost-effectiveness, velocity measurement, and privacy-preserving features, positioning it as a promising sensor-driven solution towards robust spatial perception. Yet, it remains largely unknown how to harness 4D radar for effective spatial perception, and research in this field is scarce and particularly challenging due to two key reasons. First, 4D radar sensors have only recently become commercially available, lacking mature research infrastructure like large-scale annotated datasets and toolkit compared to cameras and LiDARs. Second, 4D radar data is inherently sparse, noisy [14] and exists in various representations (point cloud, tensor, ADC samples) [10, 13], necessitating novel approaches to fully exploit its unique sensing characteristics.

Aiming to bridge the gap and unlock the full potential of 4D radar for spatial perception, my research objectives include:

- Investigating the use of 4D radar across a variety of spatial perception tasks while recognizing challenges induced by radar data characteristics and infrastructure.
- Developing novel 4D radar-based methods tailored for individual spatial perception tasks, addressing both sensor-specific challenges and task-oriented problems.
- Evaluating these methods and demonstrate that 4D radar can serve as a robust alternative to optical sensors in challenging environments for spatial perception.

My long-term research vision is to build an ecosystem for 4D radar-based spatial perception, addressing challenges in a full-stack manner and paving the way for future advancements in this field. More broadly, I aim to encourage further research, investment, and wider adoption of 4D radar technologies.

II. Contributed Research to Date

My research to date included a series of bespoke methods for 4D radar-based scene flow estimation [4, 5], moving object detection and tracking [15], and 3D occupancy prediction [7], addressing spatial perception across multiple levels.

A. 4D Radar-based Scene Flow Estimation

A crucial spatial perception task is understanding the motion of ambient dynamic objects and the ego-agent. One representation of such motion is scene flow - a set of point-wise displacement vectors that describe the 3D motion between consecutive frames relative to the ego-agent [11]. Accurate scene flow estimation enables a holistic understanding of dynamic environments and serves as a cornerstone for essential spatial perception subtasks. Therefore, I first investigate the problem of 4D radar-based scene flow estimation in [4, 5].

*1) **Self-supervised learning**:* Estimating scene flow from 4D radar presents unique challenges due to the inherent sparsity and noise in radar point clouds, as well as the lack of point-wise scene flow annotations, which are costly to acquire in real-world settings. To tackle these challenges, we present the first study on scene flow estimation using 4D radar data [4]. A self-supervised learning method called `RaFlow` is introduced to estimate scene flow on 4D radar point clouds. A novel architecture and three loss functions are specifically designed to address the challenges induced by the characteristics of radar sensors. Without the need of annotated labels, we can collectively regularize the model to learn to estimate scene flow by exploiting the underlying supervision signals embedded in the radar measurements. `RaFlow` achieves state-of-the-art performance on 4D radar scene flow estimation, and can enable downstream motion segmentation task.

*2) **Cross-modal supervised learning**:* Despite the progress, the performance of `RaFlow` is still somewhat limited due to the lack of real supervision signals, making it less reliable for safety-critical mobile autonomy scenarios. Motivated by the fact that autonomous vehicles are equipped with multiple heterogeneous sensors, *e.g.,* LiDARs, cameras and GPS/INS, we envision that this co-located sensing redundancy can be leveraged to provide cross-modal supervision cues to 4D radar scene flow learning. Based on this insight, we propose a novel cross-modal supervised approach, `CMFlow` for 4D radar scene flow learning [5]. `CMFlow` overcomes the trade-off between annotation efforts and model performance by using complementary supervision signals retrieved from co-located heterogeneous sensors. To bootstrap the cross-supervised learning, `CMFlow` applies a multi-task model architecture and subtly combine different types of supervision cues, formulating a multi-task learning problem. `CMFlow` outperforms all baseline methods, and can even surpass fully-supervised method [18] when sufficient unannotated samples are used in our training. `CMFlow` can also support two downstream tasks, *i.e.,* motion segmentation and ego-motion estimation.

Beyond mobile autonomy, we extend radar scene flow to serve as an intermediate feature of point clouds for enhancing downstream human motion sensing tasks (*e.g.,* human action recognition and parsing) [6], demonstrating its generalization across diverse radar applications.

B. Moving Object Detection and Tracking with 4D Radar

Beyond scene flow estimation, another crucial aspect of spatial perception is robustly tracking moving objects in 3D space. This capability is pivotal for subsequent autonomy tasks, such as trajectory prediction, obstacle avoidance, and path planning [9]. However, integrating 4D radars into moving object tracking presents significant challenges. For instance, when applying a LiDAR-based method directly to 4D radar data [14], performance on object-level perception degrades by about 40%, This is because that the *tracking-by-detection* paradigm struggles when adapted to 4D radar data due to the inherent radar noise and point sparsity, undermining accurate type classification and bounding box regression.

Recognizing the challenges posed by radar noise and point sparsity in 4D radar data, we introduce `RaTrack`, a pioneering and tailored solution for moving object tracking using 4D radar point clouds [15]. `RaTrack` provides a novel perspective on the tracking of moving objects, emphasizing the utility of motion segmentation and clustering over the conventional dependence on specific object types and bounding boxes. This method also leverages insights from scene flow estimation in my previous works [4, 5], inferring point-level scene flow as an explicit complement to augment the latent features of radar point clouds. This solution is architected as an end-to-end trainable network, with its training modeled as a multi-task learning endeavor. Through experiments, `RaTrack` showcases superior tracking precision of moving objects, largely surpassing the performance of the state of the arts that depend on the tracking-by-detection paradigms.

C. 3D Occupancy Prediction with 4D Imaging Radar

In recent years, 3D occupancy-based perception has gained increasing traction due to its comprehensive and open-set depiction of scene geometry [21]. Such a unified scene representation enables a richer and more generalizable spatial understanding than traditional object-centric representations as we explored in Sec. II-B, making it particularly effective for handling corner cases. Unlike object-based perception, which focuses on the foreground entities (*e.g.,* car, pedestrians), 3D occupancy prediction require reasoning all occupied spaces, encompassing both foreground and background elements such as roads and barriers. However, radar signals reflected by low-reflectivity materials, such as the surfaces of highways, are often lost during radar point cloud generation. Therefore, conventional 'LiDAR-inspired' framework, which rely on 4D radar point clouds as input for perception, struggle to achieve reliable 3D occupancy prediction.

To avoid the loss of negligible signal returns, we advocate the usage of 4D radar tensors (4DRTs) for 3D occupancy prediction instead of 4D radar point clouds. This raw data format preserves the entirety of radar measurements, addressing the shortcomings associated with the sparseness of radar point clouds caused by the signal post-processing. Building upon this insight, we develop a novel pipeline, called `RadarOcc` in [7], for 4DRT-based 3D occupancy prediction. `RadarOcc` innovatively addresses the challenges associated with the voluminous and noisy 4D radar data by employing Doppler bins descriptors, sidelobe-aware spatial sparsification, and range-wise self-attention mechanisms. To minimize the interpolation errors associated with direct coordinate transformations, we also devise a spherical-based feature encoding followed by spherical-to-Cartesian feature aggregation. The results demonstrate `RadarOcc`'s state-of-the-art performance in radar-based 3D occupancy prediction and promising results even when compared with LiDAR or camera-based methods.

III. Future Research Directions

My current research validates the ability of 4D radar as a independent modality for robust spatial perception in mobile autonomy. To further explore its potential, I plan to focus on the following key directions in my future work:

Surrounding 4D radar perception. Current 4D radar solutions typically rely on a single front-facing sensor [14, 13], resulting in a limited field of view and constrained perception performance, which restrict its capability to function as a standalone sensor rather than merely a supplement to LiDARs. While some scanning radars (*e.g.*, Navtech [2]) offer 360° coverage, they differ from the single-chip 4D radars considered here in sensing mechanisms, cost, and applications. In response, my future work will integrate multiple 4D radars on a vehicle for surrounding perception, allowing us to develop a more competitive alternative to LiDAR-based methods. Fusing data from multiple sensors not only boosts point cloud density but also allows cross-sensor validation for filtering noise. Despite the additional hardware, this multi-radar configuration remains more cost-effective than LiDAR, while significantly enhancing radar-based perception performance.

4D radar data generation. Training 4D radar perception models is challenged by inconsistent sensor viewpoints across vehicles and the scarcity of large-scale radar datasets. The former leads to mismatched training distributions, while the latter limits coverage of real-world conditions—both of which hinder model generalization. To address these issues, a promising direction is to develop generative models that synthesize novel 4D radar measurements conditioned on existing radar or LiDAR/camera data. First, generating radar data from new viewpoints can help standardize training inputs and enable real-time viewpoint adaptation at inference. Second, synthesizing 4D radar from RGB or LiDAR data can enrich training datasets, enhancing model robustness in data-scarce scenarios.

REFERENCES

[1] Claudine Badue, Rânik Guidolini, Raphael Vivacqua Carneiro, Pedro Azevedo, Vinicius B Cardoso, Avelino Forechi, Luan Jesus, Rodrigo Berriel, Thiago M Paixao, Filipe Mutz, et al. Self-driving cars: A survey. *Expert Systems with Applications*, 165:113816, 2021.

[2] Dan Barnes, Matthew Gadd, Paul Murcutt, Paul Newman, and Ingmar Posner. The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset. In *Proceedings of the IEEE international conference on robotics and automation (ICRA)*, pages 6433–6438. IEEE, 2020.

[3] Mario Bijelic, Tobias Gruber, and Werner Ritter. Benchmarking image sensors under adverse weather conditions for autonomous driving. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, pages 1773–1779. IEEE, 2018.

[4] Fangqiang Ding, Zhijun Pan, Yimin Deng, Jianning Deng, and Chris Xiaoxuan Lu. Self-supervised scene flow estimation with 4-d automotive radar. *IEEE Robotics and Automation Letters*, 7(3):8233–8240, 2022.

[5] Fangqiang Ding, Andras Palffy, Dariu M Gavrila, and Chris Xiaoxuan Lu. Hidden gems: 4d radar scene flow learning using cross-modal supervision. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9340–9349, 2023.

[6] Fangqiang Ding, Zhen Luo, Peijun Zhao, and Chris Xiaoxuan Lu. milliflow: Scene flow estimation on mmwave radar point cloud for human motion sensing. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 202–221. Springer, 2024.

[7] Fangqiang Ding, Xiangyu Wen, Yunzhou Zhu, Yiming Li, and Chris Xiaoxuan Lu. Radarocc: Robust 3d occupancy prediction with 4d imaging radar. *Advances in Neural Information Processing Systems*, 37:101589–101617, 2024.

[8] Lili Fan, Junhao Wang, Yuanmeng Chang, Yuke Li, Yutong Wang, and Dongpu Cao. 4d mmwave radar for autonomous driving perception: A comprehensive survey. *IEEE Transactions on Intelligent Vehicles*, 2024.

[9] Shuman Guo, Shichang Wang, Zhenzhong Yang, Lijun Wang, Huawei Zhang, Pengyan Guo, Yuguo Gao, and Junkai Guo. A review of deep learning-based visual multi-object tracking algorithms for autonomous driving. *Applied Sciences*, 12(21):10741, 2022.

[10] Andrew Kramer, Kyle Harlow, Christopher Williams, and Christoffer Heckman. Coloradar: The direct 3d millimeter wave radar dataset. *The International Journal of Robotics Research*, 41(4):351–360, 2022.

[11] Xingyu Liu, Charles R Qi, and Leonidas J Guibas. Flownet3d: Learning scene flow in 3d point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 529–537, 2019.

[12] Kyle C Newton, Andrew B Gill, and Stephen M Kajiura. Electroreception in marine fishes: Chondrichthyans. *Journal of Fish Biology*, 95(1):135–154, 2019.

[13] Dong-Hee Paek, Seung-Hyun Kong, and Kevin Tirta Wijaya. K-radar: 4d radar object detection for autonomous driving in various weather conditions. *Advances on Neural Information Processing Systems (NeurIPS)*, 35: 3819–3829, 2022.

[14] Andras Palffy, Ewoud Pool, Srimannarayana Baratam, Julian FP Kooij, and Dariu M Gavrila. Multi-class road user detection with 3+1d radar in the view-of-delft dataset. *IEEE Robotics and Automation Letters*, 7(2): 4961–4968, 2022.

[15] Zhijun Pan, Fangqiang Ding, Hantao Zhong, and Chris Xiaoxuan Lu. Ratrack: moving object detection and tracking with 4d radar point cloud. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 4480–4487. IEEE, 2024.

[16] Shunqiao Sun and Yimin D Zhang. 4d automotive radar sensing for autonomous vehicles: A sparsity-oriented approach. *IEEE Journal of Selected Topics in Signal Processing*, 15(4):879–891, 2021.

[17] Jessica Van Brummelen, Marie O'brien, Dominique Gruyer, and Homayoun Najjaran. Autonomous vehicle perception: The technology of today and tomorrow. *Transportation Research Part C: Emerging Technologies*, 89:384–406, 2018.

[18] Yi Wei, Ziyi Wang, Yongming Rao, Jiwen Lu, and Jie Zhou. Pv-raft: Point-voxel correlation fields for scene flow estimation of point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6954–6963, 2021.

[19] Li-Hua Wen and Kang-Hyun Jo. Deep learning-based perception systems for autonomous driving: A comprehensive survey. *Neurocomputing*, 489:255–270, 2022.

[20] Roswitha Wiltschko, Ingo Schiffner, Patrick Fuhrmann, and Wolfgang Wiltschko. The role of the magnetite-based receptors in the beak in pigeon homing. *Current Biology*, 20(17):1534–1538, 2010.

[21] Huaiyuan Xu, Junliang Chen, Shiyu Meng, Yi Wang, and Lap-Pui Chau. A survey on occupancy perception for autonomous driving: The information fusion perspective. *Information Fusion*, 114:102671, 2025.

[22] Ekim Yurtsever, Jacob Lambert, Alexander Carballo, and Kazuya Takeda. A survey of autonomous driving: Common practices and emerging technologies. *IEEE Access*, 8:58443–58469, 2020.

[23] Yuxiao Zhang, Alexander Carballo, Hanting Yang, and Kazuya Takeda. Perception and sensing for autonomous vehicles under adverse weather conditions: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing*, 196:146–177, 2023.