

Long-SCOPE: Fully Sparse Long-Range Cooperative 3D Perception

Jiahao Wang¹, Zikun Xu¹, Yuner Zhang², Zhongwei Jiang³, Chenyang Lu¹, Shuocheng Yang¹
 Yuxuan Wang¹, Jiaru Zhong⁴, Chuang Zhang¹, Shaobing Xu^{1*}, Jianqiang Wang^{1*}

¹Tsinghua University, ²University of Pennsylvania

³Nanyang Technological University, ⁴The Hong Kong Polytechnic University

wjh22@mails.tsinghua.edu.cn, {shaobxu, wjqlws}@tsinghua.edu.cn

Abstract

Cooperative 3D perception via Vehicle-to-Everything communication is a promising paradigm for enhancing autonomous driving, offering extended sensing horizons and occlusion resolution. However, the practical deployment of existing methods is hindered at long distances by two critical bottlenecks: the quadratic computational scaling of dense BEV representations and the fragility of feature association mechanisms under significant observation and alignment errors. To overcome these limitations, we introduce Long-SCOPE, a fully sparse framework designed for robust long-distance cooperative 3D perception. Our method features two novel components: a Geometry-guided Query Generation module to accurately detect small, distant objects, and a learnable Context-Aware Association module that robustly matches cooperative queries despite severe positional noise. Experiments on the V2X-Seq and Griffin datasets validate that Long-SCOPE achieves state-of-the-art performance, particularly in challenging 100–150 m long-range settings, while maintaining highly competitive computation and communication costs.

1. Introduction

Despite substantial advancements in autonomous driving technologies [10, 39, 42], current single-vehicle systems remain fundamentally constrained by limited sensor fields-of-view, performance degradation at long ranges, and severe occlusions. To overcome these individual limitations, cooperative perception has emerged as a key paradigm. By enabling information exchange between multiple agents such as vehicles (V2V), infrastructure (V2I), and drones (V2D), it promises to compensate for the deficiencies of single-vehicle systems, particularly by resolving occlusions and extending the effective perception horizon.

However, a fundamental conflict between performance

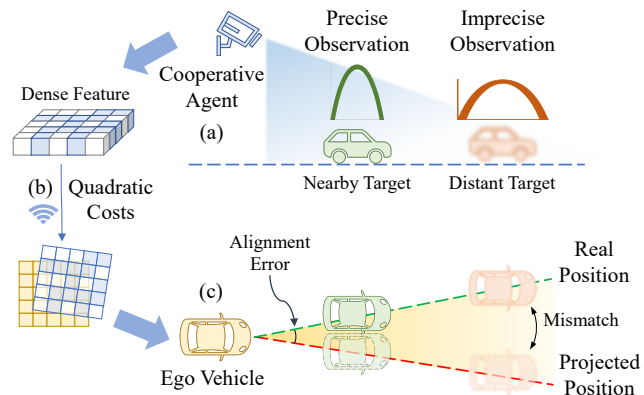


Figure 1. Core challenges in long-range cooperation: (a) observation imprecision for distant targets, (b) quadratic scaling costs of dense features, and (c) positional mismatch from alignment errors.

and cost limits the practical deployment of this vision. Effective cooperation necessitates a high density of participating agents, yet prevailing methods often depend on expensive LiDAR sensors, failing to leverage already widespread, low-cost roadside cameras. Moreover, the associated communication and computational demands impose heavy hardware infrastructure costs. Therefore, a critical bottleneck to real-world deployment is maximizing long-range perception performance using low-cost visual sensors within strictly limited communication and computational budgets. This introduces two primary challenges:

First is the choice of information structure for transmission. As shown in Fig. 1(b), mainstream cooperative perception methods rely on dense BEV features [14, 18, 23, 32], for which both computation and communication costs scale quadratically with the perception range. Although recent methods transmitting sparse feature structures have enabled communication costs to scale linearly with the number of targets in the scene, their frameworks still depend on dense BEV backbones [22, 31, 36, 47, 49], thereby inheriting the same computational scaling limitations.

Second is the difficulty of accurately fusing cooperative

*Corresponding authors

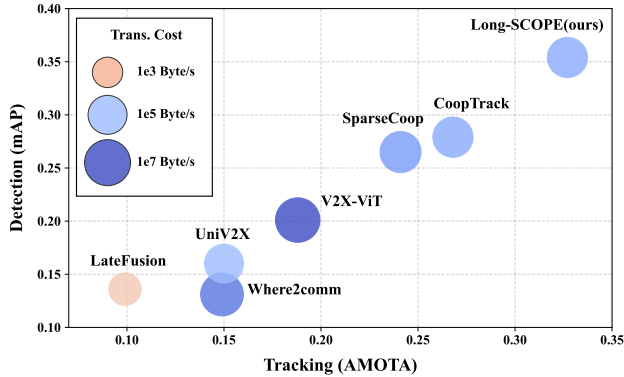


Figure 2. Performance comparison on *Griffin-25m* dataset. The X-axis and Y-axis represent perception metrics, while the bubble size and color encode the transmission cost on a logarithmic scale.

information after transmission. As shown in Fig. 1(a), for distant targets, the information extracted and sent by each agent inherently carries greater observation error. Simultaneously, calibration errors and localization noise between agents create alignment errors, leading to severe positional deviations for distant targets during coordinate projection (Fig. 1(c)). Most existing cooperative fusion methods rely on fixed distance thresholds, which are highly fragile under the influence of these two factors, leading to incorrect matches and increased ambiguity.

To address these problems, we propose **Long-SCOPE**, a **Long-Range Sparse COoperative PE**reception framework designed to push cooperative perception toward practical, long-range application. Our method completely abandons BEV features, adopting a fully sparse architecture that extracts object queries directly from image features. The main contributions are summarized as follows:

- To reduce observation error in distant targets, we introduce a **Geometry-guided Query Generation (GQG)** module. This component utilizes 2D and position priors to dynamically generate accurately located 3D queries at the single-agent level, significantly enhancing initial detection quality.
- To overcome fusion failures from alignment errors, we design a novel **Context-Aware Association (CAA)** module. It leverages both spatial context and semantic similarity of queries to achieve robust matching, resolving cooperative-level ambiguity at long distances and demonstrating superior resilience to positional noise.
- We are the first to establish comprehensive long-range visual perception benchmarks on the *V2X-Seq* [44] and *Griffin-25m* [29] datasets. We rigorously evaluate different cooperative paradigms and demonstrate that Long-SCOPE achieves state-of-the-art (SOTA) performance, setting a new standard for both accuracy and efficiency, as highlighted in Fig. 2.

2. Related Work

Cooperative perception strategies are broadly categorized into data-level (early fusion) [1, 5, 26], result-level (late fusion) [7], and feature-level (intermediate fusion). Due to the prohibitive communication costs of early fusion and the information loss from early quantization in late fusion, feature-level fusion has become the predominant trade-off. However, mainstream feature-level methods rely on dense Bird’s-Eye-View (BEV) representations [3, 4, 23, 32], imposing quadratic computational and communication scaling relative to the perception range. While compression [15, 37] and salient region selection [13, 46] reduce bandwidth, the fundamental quadratic computational bottleneck remains.

To address this, an efficient paradigm [6, 9, 48] using sparse, object-level queries has emerged, achieving communication costs that scale linearly with the number of targets rather than the perception range. However, many of these “sparse-communication” frameworks still depend on dense BEV backbones for query generation [22, 31, 36, 47, 49], reintroducing the quadratic computational bottleneck. Recently, SparseCoop [30] introduced the first fully sparse architecture, which serves as our baseline.

A critical remaining challenge in these sparse frameworks is the robust association of queries, particularly in the presence of long-range positional mismatches. Early methods [6, 22] naively scattered cooperative queries into the ego BEV grid, largely ignoring observation errors. Subsequent approaches employed explicit pairwise matching based on relative distance [8, 45, 48], feature similarity [30, 36], or position embeddings [47, 49]. To mitigate the fragility of fixed distance thresholds [8, 36], recent works [30, 45, 47, 48] employ bipartite matching via the Hungarian algorithm [17] for global optimization. However, a fundamental limitation persists: the handling of non-corresponding queries (e.g., ego-missed or invisible targets in Fig. 3). Even globally optimized methods typically resort to a final, non-adaptive distance threshold to reject unmatched queries. Under the significant positional noise inherent to long distances, this rigid heuristic inadvertently filters out valid cooperative queries. Long-SCOPE directly addresses this precise vulnerability of previous methods.

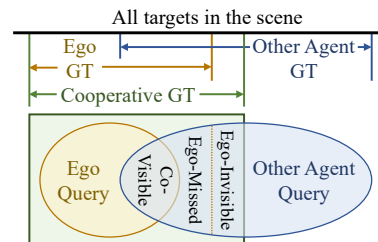


Figure 3. Task Definition of Cooperative Perception. Co-Visible, Ego-Missed, and Ego-Invisible cases should be correctly handled.

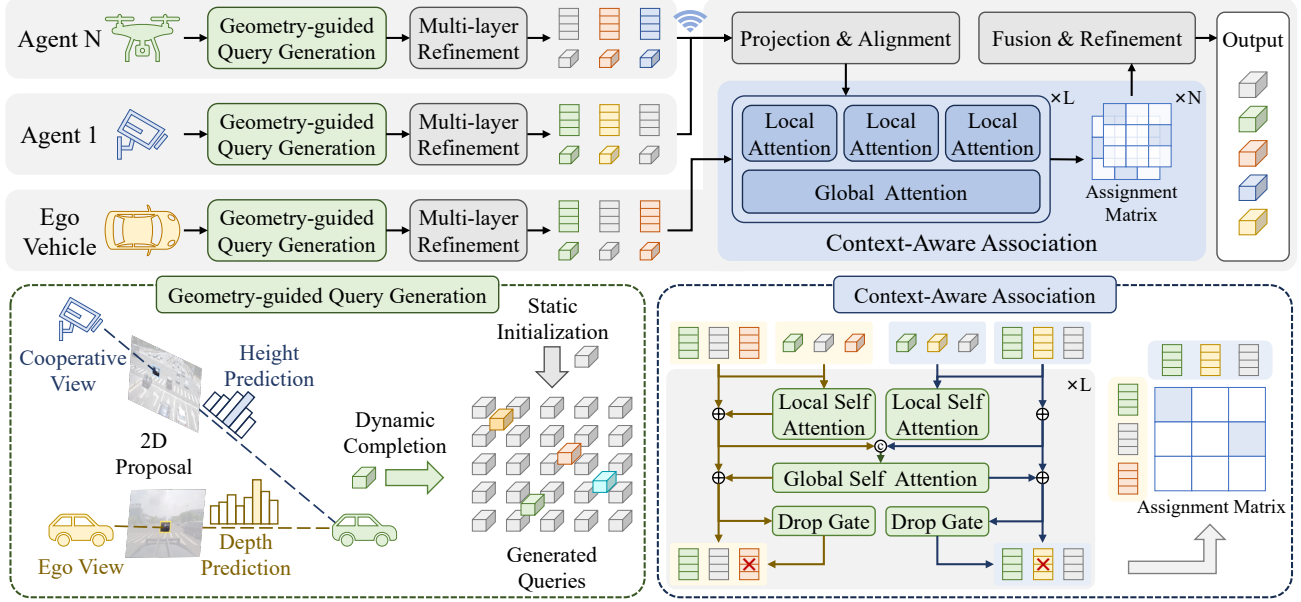


Figure 4. Overview of Long-SCOPE framework, highlighting our novel components: the Geometry-guided Query Generation module for dynamically proposing object queries and the Context-Aware Association module for robust multi-agent query matching.

3. Method

3.1. Preliminaries: Task Definition

We adopt the standard task formulation from prominent cooperative perception benchmarks [29, 43, 44]. As shown in Fig. 3 (Top), the cooperative perception ground truth GT is achieved by unifying annotations from the ego-agent GT_{ego} and cooperative agents GT_{co} , which are transformed into the ego-agent’s coordinate system, denoted as \widetilde{GT}_{co} . The resulting set is then filtered to include only objects within an ego-centered Region of Interest (ROI), R_{ego} . The final set is defined as:

$$GT = \{o \mid o \in GT_{ego} \cup \widetilde{GT}_{co}, c(o) \in R_{ego}\}$$

where o is a ground truth object and $c(o)$ is its geometric center. As shown in Fig. 3 (Bottom), achieving this GT requires not only associating co-visible queries under positional noise, but also correctly integrating ego-missed and ego-invisible queries to enrich the scene, all while avoiding the creation of duplicate detections.

3.2. Pipeline Overview

Long-SCOPE is a fully sparse, query-centric framework that circumvents the quadratic scaling costs of traditional dense BEV feature maps. As illustrated in Fig. 4, our pipeline proceeds as follows: First, at the single-agent level, each agent initiates a set of object queries. This set is a consolidated collection of both static, fixed anchors and dynamic queries generated by our novel Geometry-guided Query Generation (GQG) module. These combined

queries are then processed by a multi-layer transformer decoder [20] to refine their semantic features and state vectors. The queries from cooperative agents are then projected and aligned into the ego-vehicle’s coordinate system to compensate for spatio-temporal differences. Next, all queries—both local and aligned cooperative—are fed into our Context-Aware Association (CAA) module, which robustly matches corresponding queries and identifies unique (e.g., ego-invisible) objects. Finally, the resulting set of matched, unmatched, and local queries is processed by a fusion and refinement network to produce the comprehensive 3D detection output.

Our architecture leverages the foundational sparse paradigm of SparseCoop [30], adopting its core modules for multi-layer refinement, spatio-temporal alignment, and final fusion. However, this baseline is fragile under long-range conditions, a limitation rooted in its core association logic. To maintain performance, it relies on a simple heuristic: associating only cooperative queries within a 30m or even 15m radius of the ego-vehicle, while naively forwarding all queries beyond this range without any association. This simplistic approach suffers from duplicate detections and fails to resolve the significant observation and alignment errors inherent in long-distance perception.

To overcome these limitations, Long-SCOPE introduces two novel components that replace the baseline’s most vulnerable modules. First, to address observation errors, the GQG module (Sec. 3.3) augments static anchors with dynamically generated queries informed by geometric priors, significantly improving initial detection quality for distant

targets. Second, to resolve projection errors and fragile matching, the CAA module (Sec. 3.4) replaces rule-based association with a learnable, attention-based mechanism. By leveraging both spatial context and semantic similarity, it achieves robust query association well beyond the 30m limit, forming the core of our long-range fusion capability.

3.3. Geometry-guided Query Generation

The conventional approach of initializing object queries from a fixed set of static anchors [30, 33] is suboptimal for long-range perception. Distant objects are small and sparsely distributed, meaning most static queries will be initialized far from any valid target, compromising perception performance.

To address this, we introduce the Geometry-guided Query Generation module, which aims to dynamically generate a set of high-quality, data-driven queries by leveraging 2D image priors. Similar to Far3D [16], we first employ lightweight 2D detection and monocular depth estimation heads on the shared image features. For a target Q in the scene, these heads provide a coarse 2D proposal centered at pixel $P_{\text{img}} = [u, v, 1]^T$ and an initial depth estimate $\hat{z}_{Q_{\text{cam}}}$.

However, direct depth regression is an ill-posed problem. Inspired by BEVHeight [40, 41] and DHD [35], for agents with a high vantage point (e.g., roadside units, drones), we can simplify this task by instead predicting the object’s global height, $\hat{z}_{Q_{\text{glb}}}$. As illustrated by the statistics in Fig. 5, the object height distribution in typical driving scenes is highly concentrated, whereas the depth distribution is extremely dispersed. This makes global height a significantly more stable and well-posed target for optimization.

Consequently, for high-vantage agents, we re-frame the depth estimation problem as a geometric derivation from this more stable height prediction. This derivation requires three coordinate frames: (1) the global frame \mathcal{F}_{glb} (e.g., East-North-Up); (2) the camera frame \mathcal{F}_{cam} , with its origin C at position C_{glb} in \mathcal{F}_{glb} , where the projection matrix $T_{\text{cam2glb}} \in \mathbb{R}^{4 \times 4}$ is known from calibration and localization data; and (3) a “virtual” frame $\mathcal{F}_{\text{virt}}$, also centered at C but with axes parallel to \mathcal{F}_{glb} , sharing the rotation $R_{\text{cam2virt}} \in \mathbb{R}^{3 \times 3}$ from $T_{\text{cam2glb}}[:, :3]$.

Our 2D-to-3D lifting for high-vantage agents, diagrammed in Fig. 5, begins with a network head predicting the target’s global height, $\hat{z}_{Q_{\text{glb}}}$, for a 2D proposal at pixel P_{img} . First, we compute the target’s height in the virtual frame, $\hat{z}_{Q_{\text{virt}}} = \hat{z}_{Q_{\text{glb}}} - z_{C_{\text{glb}}}$, by subtracting the camera’s known global height $z_{C_{\text{glb}}}$. Next, we find the corresponding ray’s virtual height by un-projecting the pixel P_{img} to a 3D ray in the camera frame, $P_{\text{cam}} = K_{\text{cam}}^{-1} \cdot P_{\text{img}}$, assuming unit depth ($z_{P_{\text{cam}}} := 1$). This ray is transformed into the virtual frame via $P_{\text{virt}} = R_{\text{cam2virt}} \cdot P_{\text{cam}}$, and we extract its z -component, $z_{P_{\text{virt}}}$. As shown in Fig. 5, the geometry forms two similar triangles, allowing us to equate the ratio of the

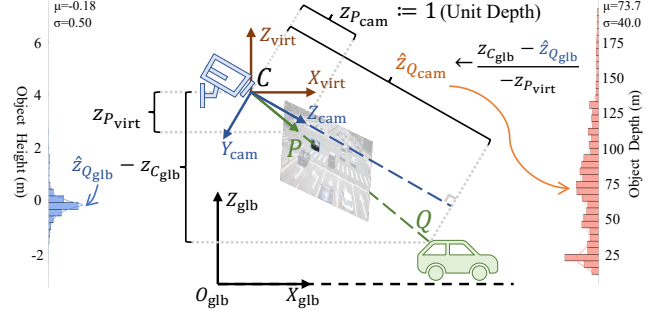


Figure 5. Depth derivation for high-vantage agents. We predict the stable global height $\hat{z}_{Q_{\text{glb}}}$ (left histogram) instead of the dispersed depth $\hat{z}_{Q_{\text{cam}}}$ (right histogram) and utilize geometric priors.

target’s camera depth $\hat{z}_{Q_{\text{cam}}}$ to the unit depth (1) with the ratio of the target’s virtual height $\hat{z}_{Q_{\text{virt}}}$ to the ray’s virtual height $z_{P_{\text{virt}}}$. We can therefore solve for the unknown depth $\hat{z}_{Q_{\text{cam}}}$ directly:

$$\hat{z}_{Q_{\text{cam}}} = \frac{\hat{z}_{Q_{\text{virt}}}}{z_{P_{\text{virt}}}} = \frac{\hat{z}_{Q_{\text{glb}}} - z_{C_{\text{glb}}}}{(T_{\text{cam2glb}}[:, :3] \cdot K_{\text{cam}}^{-1} \cdot P_{\text{img}})_z} \quad (1)$$

This derivation provides a stable depth estimate. For high-angle views, $z_{P_{\text{virt}}}$ is significantly non-zero, avoiding numerical issues.

Conversely, for ground-level agents (e.g., vehicles), the camera view is nearly parallel to the ground plane. For pixels near the horizon, $z_{P_{\text{virt}}} \approx 0$, making the division in Eq. (1) numerically unstable. Therefore, for these agents, we revert to the standard approach and use a lightweight head \mathcal{D} to directly regress the depth $\hat{z}_{Q_{\text{cam}}} = \mathcal{D}(f_{\text{img}}(P_{\text{img}}))$ using the image features $f_{\text{img}}(P_{\text{img}})$.

Regardless of which estimation strategy is employed, we obtain a 3D position P_{cam} in the camera frame \mathcal{F}_{cam} by un-projecting the 2D proposal P_{img} with its estimated depth $\hat{z}_{Q_{\text{cam}}}$ as $P_{\text{cam}} = \hat{z}_{Q_{\text{cam}}} \cdot (K_{\text{cam}}^{-1} P_{\text{img}})$. This point is then transformed into the agent’s local coordinate system using known camera extrinsics. This 3D position, along with the corresponding image feature extracted via MaxPooling, are used to initialize a new dynamic query. These dynamically generated queries are then merged with the set of static, fixed-initialized queries and fed collectively into the multi-layer 3D decoder, significantly improving its ability to find and refine distant targets.

3.4. Context-Aware Association

The core bottleneck in long-range cooperative fusion lies in the robust association of sparse queries. As demonstrated by the baseline limitations in Sec. 3.2, simple heuristics such as distance-based cutoffs are highly sensitive to the significant observation and projection noise inherent to distant targets. This fragility results in duplicate detections from unassociated queries and a failure to capitalize on the

extended field-of-view that cooperation aims to deliver. To overcome this, we introduce the Context-Aware Association module, which replaces rigid heuristics with a learnable, context-driven matching mechanism.

To ensure robustness against long-range ambiguities, its architecture is governed by four core design principles:

- **Unambiguous Matching:** A single target should generate at most one query per agent. Therefore, valid associations should strictly enforce injective (one-to-one) correspondences, preventing a single query from being matched multiple times.
- **Asymmetric Visibility:** The formulation must natively support unmatchable queries arising from ego-missed or ego-invisible targets, seamlessly preserving them for the final perception output.
- **Spatial Consistency:** While the absolute coordinates of distant targets are frequently corrupted by alignment errors, their local spatial context remains a highly stable signature. A successful match must demonstrate consistency not only in semantic features but also within its local neighborhood topology.
- **Scalability:** The association framework must generalize efficiently to a dynamic number of N cooperative agents, extending beyond simple pairwise matching.

From this perspective, existing rule-based approaches with fixed distance thresholds [8, 36] satisfy only the Asymmetric Visibility requirement. While set-based bipartite matching [30, 45, 48] resolves Unambiguous Matching, it still ignores Spatial Consistency and is generally confined to pairwise fusion without Scalability.

To comprehensively satisfy these principles, our CAA module employs a multi-layer Transformer [27] architecture inspired by recent advancements in robust feature matching [21, 24]. Unlike standard cross-attention mechanisms restricted to two sets, we further adopt an extensible Structure-from-Motion (SfM) approach [28, 34]: we concatenate the queries from all N agents and utilize global self-attention. While this incurs a quadratic computational cost relative to the query count, the total number of high-confidence transmitted queries is tied to the physical targets in the scene, generally remaining below 100 and ensuring high efficiency.

Contextual Feature Refinement. As shown in Fig. 4, the CAA module refines query descriptors iteratively over L layers. Each layer consists of two sequential operations: First, to encode *Spatial Consistency*, an intra-agent (local) self-attention block operates independently on each agent’s query set. This block utilizes positional encodings derived from the queries’ state vectors, allowing each query to aggregate topological information from its spatial neighbors. This enriches the descriptor with a stable relative context of its surrounding targets. Second, an inter-agent (global) self-attention block operates across the concatenated pool

of all N agents’ queries. Crucially, positional encoding is disabled here to achieve invariance to projection noise. This unconstrained global attention enables queries representing the same physical object—now augmented with robust local topology—to implicitly identify correspondences and mutually align their semantic representations.

Optimal Transport Matching. After L refinement layers, the module produces highly contextualized query descriptors. To extract discrete associations, we decouple the global multi-way matching into parallelized pairwise assignment heads: for each cooperative agent i , we compute an affinity matrix S between its refined queries $Q_{coop,i}$ and the ego queries Q_{ego} using feature inner products. Following SuperGlue [24], we apply Sinkhorn normalization [25] to S to derive a doubly stochastic partial assignment matrix $P(u, x)$, representing the matching probability between cooperative query u and ego query x . To yield the final assignment matrix, we scale these probabilities by per-descriptor matchability scores (σ^u, σ^x) as formulated in LightGlue [21]. The resulting matches are filtered via a mutual nearest-neighbor check and a fixed confidence threshold. Matched queries proceed to the fusion and refinement stage, while the unmatched cooperative queries, representing uniquely observed ego-missed or ego-invisible targets, are directly appended to the final output.

4. Experiments

4.1. Datasets and Metrics

V2X-Seq. This is a large-scale, sequential dataset designed for V2I cooperative 3D object detection and tracking [44]. It features sensor data from both an ego-vehicle and a connected RSU, capturing complex urban traffic scenarios with significant occlusions. We adhere to the official data splits and protocol of the CVPR 2025 challenge [11]. Critically, to rigorously benchmark the long-range performance central to our work, we expand the evaluation scope from the standard 50 meters to 150 meters.

Griffin. This is a pioneering simulated dataset for aerial-ground cooperative perception [29]. It presents unique challenges due to the large viewpoint disparity and dynamic transformations inherent to V2D scenarios. We again follow the official protocol and data splits for evaluation. Consistent with our long-range focus, we expand the evaluation horizon from the original 50 meters to 100 meters.

Evaluation Metrics. Both benchmarks adopt established metrics from the NuScenes benchmark [2] and focus on the ‘car’ class. We primarily report Average Precision (AP) to assess detection quality and Average Multi-Object Tracking Accuracy (AMOTA) for tracking performance. To validate our claims of efficiency and practicality, we also measure transmission cost in Bytes Per Second (BPS) and computational efficiency via inference Frames Per Second (FPS).

Table 1. Performance comparison on *Griffin-25m*. **Bold**, underline, and *italic* represent the top three among feature-level fusion methods.

Method	0–100 m Overall		0–50 m		50–100 m		BPS ↓	FPS ↑
	AP ↑	AMOTA ↑	AP ↑	AMOTA ↑	AP ↑	AMOTA ↑		
Late Fusion Baseline	0.136	0.099	0.357	0.325	0.004	0.000	2.55×10^3	6.13
Where2Comm [13] (NIPS 2022)	0.131	0.149	0.384	0.438	0.000	0.000	1.05×10^6	7.34
V2X-ViT [38] (ECCV 2022)	0.201	0.188	<i>0.494</i>	<i>0.527</i>	0.011	0.000	3.21×10^6	<i>7.50</i>
UniV2X [45] (AAAI 2025)	0.160	0.150	0.417	0.428	0.007	0.000	9.84×10^4	6.51
CoopTrack [49] (ICCV 2025)	<u>0.279</u>	<u>0.268</u>	<u>0.560</u>	0.617	<i>0.054</i>	0.000	2.00×10^5	6.03
SparseCoop [30] (AAAI 2026)	<u>0.265</u>	<i>0.241</i>	<i>0.494</i>	0.478	<u>0.088</u>	<u>0.016</u>	3.31×10^5	8.87
Long-SCOPE (Ours)	0.354	0.327	0.597	<u>0.566</u>	0.151	0.112	1.90×10^5	<u>7.68</u>
Long-SCOPE w/o GQG	0.284	0.232	0.523	0.480	0.083	0.011	3.31×10^5	8.05
Long-SCOPE w/o CAA	0.329	0.294	0.542	0.495	0.155	0.131	1.90×10^5	8.10

Table 2. Performance comparison on *V2X-Seq*. **Bold**, underline, and *italic* represent the top three among feature-level fusion methods.

Method	0–150 m Overall		0–50 m		50–100 m		100–150 m	
	AP ↑	AMOTA ↑	AP ↑	AMOTA ↑	AP ↑	AMOTA ↑	AP ↑	AMOTA ↑
Late Fusion Baseline	0.224	0.271	0.318	0.375	0.178	0.185	0.050	0.032
Where2Comm [13] (NIPS 2022)	0.059	0.004	0.141	0.080	0.026	0.000	0.000	0.000
V2X-ViT [38] (ECCV 2022)	<i>0.289</i>	<u>0.345</u>	<i>0.421</i>	<i>0.509</i>	<u>0.231</u>	<u>0.251</u>	<i>0.054</i>	<u>0.027</u>
UniV2X [45] (AAAI 2025)	0.166	0.019	0.255	0.101	0.123	0.000	0.032	0.000
CoopTrack [49] (ICCV 2025)	0.232	0.171	0.327	0.255	0.192	0.088	<u>0.059</u>	<i>0.006</i>
SparseCoop [30] (AAAI 2026)	<u>0.334</u>	<i>0.328</i>	<u>0.515</u>	<u>0.534</u>	<i>0.229</i>	<i>0.232</i>	0.031	0.000
Long-SCOPE (Ours)	0.399	0.444	0.532	0.576	0.339	0.361	0.113	0.059
Long-SCOPE w/o GQG	0.388	0.416	0.524	0.559	0.321	0.341	0.113	0.046
Long-SCOPE w/o CAA	0.352	0.382	0.530	0.588	0.243	0.280	0.041	0.005

4.2. Implementation Details

Model and Baselines. All models are built upon a ResNet50 [12] backbone, pre-trained on ImageNet. For a fair comparison, baselines were reproduced as follows: UniV2X [45], CoopTrack [49], and SparseCoop [30] were implemented using their official codebases. Where2Comm [13] and V2X-ViT [38] were re-implemented using a BEVFormer [19] backbone within the shared UniV2X/CoopTrack codebase. For these BEV-based methods, we set the grid resolution to $1m \times 1m$, with the total grid size scaling proportionally to the perception range. The query-based methods [30, 45, 49] all use 900 queries. Our Long-SCOPE framework adopts the same set, which is then augmented by our novel GQG module that generates at most 40 additional dynamic queries. Our CAA module is implemented with 4 iterative refinement layers.

Training Regimen. All methods adopt the two-stage training strategy. Initially, the single-agent models are trained for 48 epochs. Subsequently, the cooperative model is initialized with the single-agent weights and fine-tuned for an

additional 24 epochs. We employ the AdamW optimizer with a total batch size of 16 and a weight decay of 0.01. The initial learning rate is set to 2×10^{-4} and is adjusted using a cosine annealing schedule.

Data and Hardware. Input images from V2X-Seq are resized from 1920×1080 to 960×540 , while images from Griffin-25m are resized to 800×448 . All experiments were performed distributed across four NVIDIA 3090 GPUs.

4.3. State-of-the-Art Comparison

As detailed in Tabs. 1 and 2, Long-SCOPE achieves state-of-the-art performance on both the aerial-ground *Griffin-25m* and real-world V2I *V2X-Seq* datasets, significantly outperforming all baselines in both detection and tracking. On *Griffin-25m*, Long-SCOPE achieves an overall AP of 0.354, outperforming the next-best feature-level method (CoopTrack) by 7.5 AP points and the sparse baseline (SparseCoop) by 8.9 AP points. Similarly, in the 0–150 m overall evaluation on *V2X-Seq*, Long-SCOPE reaches 0.399 AP and 0.444 AMOTA, surpassing SparseCoop by 6.5 AP points and the dense V2X-ViT by 11.0 AP points.

Table 3. Detailed ablation study for the GQG module. We report the performance gains of single-agent models on *Griffin-25m* dataset.

Agent Type	Query Method	0–100 m Overall		0–50 m		50–100 m	
		AP ↑	AMOTA ↑	AP ↑	AMOTA ↑	AP ↑	AMOTA ↑
Drone (High-Vantage)	Baseline (Static Queries)	0.371	0.393	0.639	0.766	0.092	0.010
	+ Dynamic (Direct Depth)	0.390	0.403	0.647	0.745	0.116	0.032
	+ Dynamic (Height-Derived)	0.449	0.443	0.694	0.748	0.182	0.099
Vehicle (Ground-Level)	Baseline (Static Queries)	0.413	0.252	0.604	0.384	0.047	0.019
	+ Dynamic (Direct Depth)	0.428	0.256	0.610	0.407	0.052	0.000

Crucially, Long-SCOPE’s advantage widens significantly as the perception range extends, proving its unique capability for true long-range fusion. In the challenging 50–100 m bracket on *Griffin-25m*, Long-SCOPE maintains robust tracking (0.112 AMOTA) and achieves 0.151 AP, demonstrating graceful degradation. In contrast, methods like CoopTrack and V2X-ViT collapse to near-zero tracking capabilities. On *V2X-Seq*, this resilience is even more pronounced: in the extreme 100–150 m bracket, Long-SCOPE (0.113 AP) nearly doubles the performance of the closest competitor, CoopTrack (0.059 AP), and is 3.6 times more accurate than SparseCoop (0.031 AP). These results confirm our framework’s effectiveness in resolving the severe positional noise and sparse observations inherent to extended distances.

System Efficiency. Beyond accuracy, Long-SCOPE shows high practical viability for real-world deployment. It maintains a highly competitive transmission cost of 1.90×10^5 BPS. This is approximately 17 times more bandwidth-efficient than the dense BEV-based V2X-ViT (3.21×10^6 BPS) and remains on par with other sparse methods, delivering markedly superior perception without inflating communication overhead. Furthermore, in terms of computational efficiency, Long-SCOPE achieves an inference speed of 7.68 FPS. This is fully competitive with both dense and sparse baselines (e.g., V2X-ViT at 7.50 FPS and UniV2X at 6.51 FPS), confirming that the substantial gains in long-range accuracy and robust query association do not come at the expense of real-time processing capabilities.

4.4. Ablation Studies

To verify the individual contributions of our novel components, we conduct extensive ablation studies, with the results summarized in Tabs. 1 to 3.

Impact of Core Components. We first validate the contributions of our two primary modules by removing them individually. The Context-Aware Association (CAA) module proves to be the cornerstone for long-range fusion. As shown in Tab. 1 and Tab. 2, omitting the CAA module causes a severe performance drop, with the overall AP falling from 0.354 to 0.329 on *Griffin-25m*, and from 0.399

to 0.352 on *V2X-Seq*. Critically, this degradation is most severe precisely where Long-SCOPE claims its largest advantage: in the 100–150 m range, AP on *V2X-Seq* plummets from 0.113 to 0.041. This sharp decline, which reduces our model’s performance to the level of competing baselines, provides direct evidence for our central thesis: the simple, heuristic-based matching used by existing frameworks is the primary failure point under long-range alignment errors. Our context-driven, learnable CAA module is the specific mechanism required to resolve this bottleneck.

Concurrently, the Geometry-guided Query Generation (GQG) module provides a foundational performance lift. Removing it consistently degrades overall performance across both datasets (0.354 to 0.284 AP on *Griffin-25m*, and 0.399 to 0.388 AP on *V2X-Seq*). This confirms that generating high-quality, data-driven queries is essential for initiating the perception pipeline with accurate proposals for distant targets, which subsequently enhances the entire multi-agent fusion process.

Validation of GQG Module Design. We further validate the GQG module in Tab. 3. For high-vantage agents (e.g., drones), our proposed height-derived depth estimation is vastly superior to direct regression. It boosts the 50–100 m AP from 0.116 (with Direct Depth) to 0.182, nearly doubling the long-range AP of the static baseline (0.092). This validates the stability of predicting global height rather than depth. Conversely, for ground-level agents (e.g., vehicles) where this geometric derivation is unstable, using standard direct depth regression for dynamic queries still provides a consistent, modest improvement over static queries alone (0.052 vs. 0.047 AP in the 50–100 m range). This ablation confirms our two-pronged strategy within the GQG module is well-founded and optimal for diverse agent types.

4.5. Robustness to Localization Noise

Real-world V2X systems are inherently susceptible to GPS and calibration deviations. To validate the practical resilience of the CAA module against these alignment errors, we investigate Long-SCOPE’s performance under increasing levels of injected localization noise in Fig. 6. We observe a distinct trade-off regarding the matching confidence

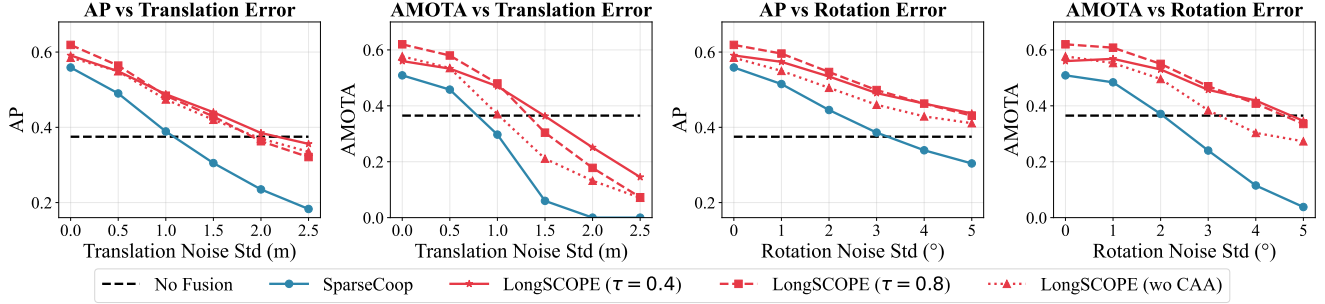


Figure 6. Comparison of robustness to localization errors on *Griffin-25m* dataset. Long-SCOPE maintains stable tracking and detection under increasing translation and rotation noise compared to the baseline.

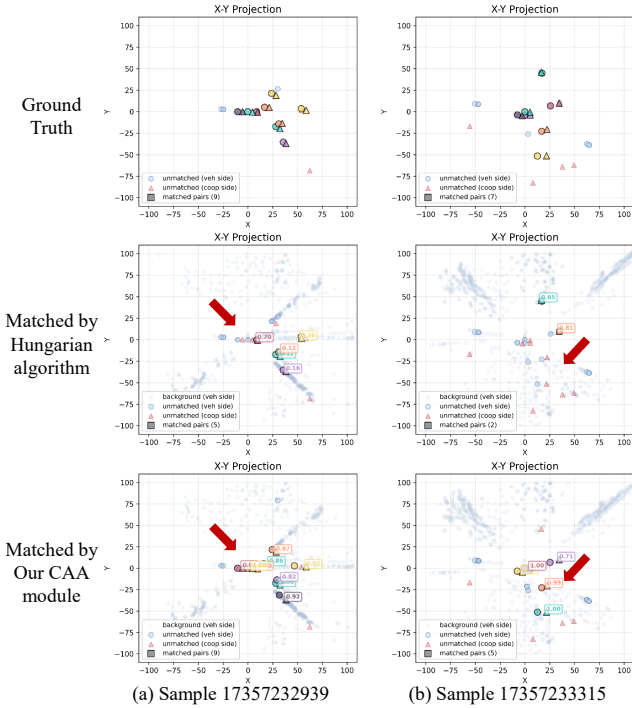


Figure 7. Qualitative comparison of query association performance under severe localization noise on *Griffin-25m* dataset. Matched query pairs share the same color. While the Hungarian algorithm (middle, red arrows) struggles with alignment errors by prioritizing geometric closeness, our CAA module (bottom) robustly resolves these ambiguities.

threshold τ . While a stricter threshold ($\tau = 0.8$) provides peak performance under ideal, noise-free localization, its criteria become overly conservative as noise increases, inadvertently filtering out some valid cooperative queries. Conversely, a more balanced threshold ($\tau = 0.4$) demonstrates superior real-world robustness, maintaining reliable association capabilities even under significant alignment error, including translation noise with a standard deviation exceeding 1.0 m or rotation noise beyond 4° . Furthermore,

Long-SCOPE equipped with the CAA module clearly outperforms both the baseline variant using standard Hungarian matching (denoted as *w/o CAA*) and the original SparseCoop [30] architecture.

Figure 7 visualizes the critical failure modes of standard association techniques compared to our proposed CAA module. Standard bipartite matching (e.g., the Hungarian algorithm [17]) relies almost exclusively on Euclidean distance. As illustrated in the middle row of Fig. 7, when significant localization noise is introduced, the projected position of a cooperative agent’s query drifts far from the ego query. Consequently, the Hungarian algorithm forces matches based on geometric proximity, often incorrectly associating distinct objects or completely failing to match corresponding ones. The CAA module overcomes this limitation by aggregating local neighborhood information via self-attention prior to the matching step. Considering that a cluster of vehicles maintains its relative shape and semantic signature even when absolute coordinates are globally shifted, this preservation of structural integrity allows Long-SCOPE to correctly associate queries despite severe coordinate misalignment.

5. Conclusion

We introduced Long-SCOPE, a fully sparse framework for long-range cooperative 3D perception that avoids the quadratic scaling costs of BEV-based methods. Our approach features two novel components: Geometry-guided Query Generation to accurately detect distant targets, and a learnable Context-Aware Association module that replaces fragile, heuristic-based matching. Experiments on *V2X-Seq* and *Griffin-25m* show Long-SCOPE achieves state-of-the-art performance, with its advantage widening dramatically at extreme distances (100–150 m) where existing methods collapse. This work validates that a fully sparse design, centered on robust, context-aware association, is essential for practical long-range cooperative perception.

Acknowledgments

This work was supported by the National Natural Science Foundation of China for the Science Fund for Creative Research Groups (No. 52221005), the Key Project (No. 52131201), and the Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (No. JYB2025XDXM123).

References

- [1] Eduardo Arnold, Mehrdad Dianati, Robert De Temple, and Saber Fallah. Cooperative perception for 3D object detection in driving scenarios using infrastructure sensors. *IEEE Transactions on Intelligent Transportation Systems*, 23(3): 1852–1864, 2022. [2](#)
- [2] Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liang, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuScenes: A multimodal dataset for autonomous driving. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11618–11628, 2020. [5](#)
- [3] Gong Chen, Chaokun Zhang, and Pengcheng Lv. CoopDiff: A Diffusion-Guided Approach for Cooperation under Corruptions, 2026. [2](#)
- [4] Gong Chen, Chaokun Zhang, Tao Tang, Pengcheng Lv, Feng Li, and Xin Xie. CATNet: Collaborative Alignment and Transformation Network for Cooperative Perception, 2026. [2](#)
- [5] Qi Chen, Sihai Tang, Qing Yang, and Song Fu. Cooper: Cooperative Perception for Connected Autonomous Vehicles Based on 3D Point Clouds. In *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*, pages 514–524, 2019. [2](#)
- [6] Ziming Chen, Yifeng Shi, and Jinrang Jia. TransIFF: An instance-level feature fusion framework for vehicle-infrastructure cooperative 3D detection with transformers. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 18159–18168, Paris, France, 2023. IEEE. [2](#)
- [7] Hsu-Kuang Chiu, Chien-Yi Wang, Min-Hung Chen, and Stephen F. Smith. Probabilistic 3D Multi-Object Cooperative Tracking for Autonomous Driving via Differentiable Multi-Sensor Kalman Filter. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 18458–18464, 2024. [2](#)
- [8] Zihan Ding, Jiahui Fu, Si Liu, Hongyu Li, Siheng Chen, Hongsheng Li, Shifeng Zhang, and Xu Zhou. Point Cluster: A Compact Message Unit for Communication-Efficient Collaborative Perception. In *The Thirteenth International Conference on Learning Representations*, 2024. [2](#), [5](#)
- [9] Siqi Fan, Haibao Yu, Wenxian Yang, Jirui Yuan, and Zaiqing Nie. QUEST: Query Stream for Practical Cooperative Perception. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 18436–18442, 2024. [2](#)
- [10] Zeyu Han, Shuocheng Yang, Minghan Zhu, Fang Zhang, Shaobing Xu, Maani Ghaffari, and Jianqiang Wang. EquIRO: A 4D mmWave Radar Odometry via Equivariant Networks. *IEEE Robotics and Automation Letters*, 11(4):4034–4041, 2026. [1](#)
- [11] Ruiyang Hao, Haibao Yu, Jiaru Zhong, Chuanye Wang, Jiahao Wang, Yiming Kan, Wenxian Yang, Siqi Fan, Huilin Yin, Jianing Qiu, Yao Mu, Jiankai Sun, Li Chen, Walter Zimmer, Dandan Zhang, Shanghang Zhang, Mac Schwager, Ping Luo, and Zaiqing Nie. Research Challenges and Progress in the End-to-End V2X Cooperative Autonomous Driving Competition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1828–1839, 2025. [5](#)
- [12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016. [6](#)
- [13] Yue Hu, Shaoheng Fang, Zixing Lei, Yiqi Zhong, and Siheng Chen. Where2comm: Communication-efficient collaborative perception via spatial confidence maps. In *Advances in Neural Information Processing Systems*, pages 4874–4886, 2022. [2](#), [6](#)
- [14] Yue Hu, Yifan Lu, Runsheng Xu, Weidi Xie, Siheng Chen, and Yanfeng Wang. Collaboration Helps Camera Overtake LiDAR in 3D Detection. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9243–9252, 2023. [1](#)
- [15] Yue Hu, Juntong Peng, Sifei Liu, Junhao Ge, Si Liu, and Siheng Chen. Communication-efficient collaborative perception via information filling with codebook. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 15481–15490, 2024. [2](#)
- [16] Xiaohui Jiang, Shuailin Li, Yingfei Liu, Shihao Wang, Fan Jia, Tiancai Wang, Lijin Han, and Xiangyu Zhang. Far3D: Expanding the Horizon for Surround-View 3D Object Detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(3):2561–2569, 2024. [4](#)
- [17] H. W. Kuhn. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97, 1955. [2](#), [8](#)
- [18] Yiming Li, Shunli Ren, Pengxiang Wu, Siheng Chen, Chen Feng, and Wenjun Zhang. Learning Distilled Collaboration Graph for Multi-Agent Perception. In *Advances in Neural Information Processing Systems*, pages 29541–29552. Curran Associates, Inc., 2021. [1](#)
- [19] Zhiqi Li, Wenhai Wang, Hongyang Li, Enze Xie, Chonghao Sima, Tong Lu, Yu Qiao, and Jifeng Dai. BEVFormer: Learning bird’s-eye-view representation from multi-camera images via spatiotemporal transformers. In *Lecture Notes in Computer Science*, pages 1–18, Cham, 2022. Springer Nature Switzerland. [6](#)
- [20] Xuwu Lin, Zixiang Pei, Tianwei Lin, Lichao Huang, and Zhizhong Su. Sparse4D v3: Advancing end-to-end 3D detection and tracking, 2023. [3](#)
- [21] Philipp Lindenberger, Paul-Edouard Sarlin, and Marc Pollefeys. LightGlue: Local Feature Matching at Light Speed. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 17581–17592, 2023. [5](#)
- [22] Haizhuang Liu, Huazhen Chu, Junbao Zhuo, Bochao Zou, Jiansheng Chen, and Huimin Ma. SparseComm: An

- Efficient Sparse Communication Framework for Vehicle-Infrastructure Cooperative 3D Detection. *Pattern Recognition*, 158:110961, 2025. 1, 2
- [23] Yifan Lu, Quanhao Li, Baoan Liu, Mehrdad Dianati, Chen Feng, Siheng Chen, and Yanfeng Wang. Robust collaborative 3D object detection in presence of pose errors. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4812–4818, 2023. 1, 2
- [24] Paul-Edouard Sarlin, Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. SuperGlue: Learning Feature Matching With Graph Neural Networks. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4937–4946, 2020. 5
- [25] Richard Sinkhorn and Paul Knopp. Concerning nonnegative matrices and doubly stochastic matrices. *Pacific Journal of Mathematics*, 21(2):343–348, 1967. 5
- [26] Rodolfo Valiente, Mahdi Zaman, Sedat Ozer, and Yaser P. Fallah. Controlling Steering Angle for Cooperative Self-driving Vehicles utilizing CNN and LSTM-based Deep Networks. In *2019 IEEE Intelligent Vehicles Symposium (IV)*, pages 2423–2428, 2019. 2
- [27] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2017. 5
- [28] Jianyuan Wang, Minghao Chen, Nikita Karaev, Andrea Vedaldi, Christian Rupprecht, and David Novotny. VGGT: Visual Geometry Grounded Transformer, 2025. 5
- [29] Jiahao Wang, Xiangyu Cao, Jiaru Zhong, Yuner Zhang, Zeyu Han, Haibao Yu, Chuang Zhang, Lei He, Shaobing Xu, and Jianqiang Wang. Griffin: Aerial-Ground Cooperative Detection and Tracking Dataset and Benchmark. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(12):9867–9875, 2026. 2, 3, 5
- [30] Jiahao Wang, Zhongwei Jiang, Wenchao Sun, Jiaru Zhong, Haibao Yu, Yuner Zhang, Chenyang Lu, Chuang Zhang, Lei He, Shaobing Xu, and Jianqiang Wang. SparseCoop: Cooperative Perception with Kinematic-Grounded Queries. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(12):9876–9884, 2026. 2, 3, 4, 5, 6, 8
- [31] Shaohong Wang, Lu Bin, Xinyu Xiao, Zhiyu Xiang, Hangguan Shan, and Eryun Liu. IFTR: An Instance-Level Fusion Transformer for Visual Collaborative Perception. In *Computer Vision – ECCV 2024*, pages 124–141, Cham, 2025. Springer Nature Switzerland. 1, 2
- [32] Tsun-Hsuan Wang, Sivabalan Manivasagam, Ming Liang, Bin Yang, Wenyan Zeng, and Raquel Urtasun. V2VNet: Vehicle-to-vehicle communication for joint perception and prediction. In *Computer Vision – ECCV 2020*, pages 605–621, Cham, 2020. Springer International Publishing. 1, 2
- [33] Yue Wang, Vitor Campagnolo Guizilini, Tianyuan Zhang, Yilun Wang, Hang Zhao, and Justin Solomon. DETR3D: 3D object detection from multi-view images via 3D-to-2D queries. In *Proceedings of the 5th Conference on Robot Learning*, pages 180–191. PMLR, 2022. 4
- [34] Yifan Wang, Jianjun Zhou, Haoyi Zhu, Wenzheng Chang, Yang Zhou, Zizun Li, Junyi Chen, Jiangmiao Pang, Chunchua Shen, and Tong He. π 3: Permutation-Equivariant Visual Geometry Learning, 2025. 5
- [35] Zhechao Wang, Peirui Cheng, Mingxin Chen, Pengju Tian, Zhirui Wang, Xinming Li, Xue Yang, and Xian Sun. Drones Help Drones: A Collaborative Framework for Multi-Drone Object Trajectory Prediction and Beyond. *Advances in Neural Information Processing Systems*, 37:64604–64628, 2024. 4
- [36] Zhe Wang, Shaocong Xu, Xucui Zhuang, Tongda Xu, Yan Wang, Jingjing Liu, Yilun Chen, and Ya-Qin Zhang. CoopDETR: A Unified Cooperative Perception Framework for 3D Detection via Object Query. In *IEEE International Conference on Robotics and Automation, ICRA 2025, Atlanta, GA, USA, May 19-23, 2025*, pages 2732–2739. IEEE, 2025. 1, 2, 5
- [37] Quanmin Wei, Penglin Dai, Wei Li, Bingyi Liu, and Xiao Wu. InfoCom: Kilobyte-Scale Communication-Efficient Collaborative Perception with Information Bottleneck. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(35):29731–29739, 2026. 2
- [38] Runsheng Xu, Hao Xiang, Zhengzhong Tu, Xin Xia, Ming-Hsuan Yang, and Jiaqi Ma. V2X-ViT: Vehicle-to-everything cooperative perception with vision transformer. In *Lecture Notes in Computer Science*, pages 107–124, Cham, 2022. Springer Nature Switzerland. 6
- [39] Zikun Xu and Shaobing Xu. MergeOcc: Bridge the Domain Gap between Different LiDARs for Robust Occupancy Prediction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 26539–26548, 2025. 1
- [40] Lei Yang, Kaicheng Yu, Tao Tang, Jun Li, Kun Yuan, Li Wang, Xinyu Zhang, and Peng Chen. BEVHeight: A robust framework for vision-based roadside 3D object detection. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 21611–21620, 2023. 4
- [41] Lei Yang, Tao Tang, Jun Li, Kun Yuan, Kai Wu, Peng Chen, Li Wang, Yi Huang, Lei Li, Xinyu Zhang, and Kaicheng Yu. BEVHeight++: Toward Robust Visual Centric 3D Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–18, 2025. 4
- [42] Shuocheng Yang, Yueming Cao, Shengbo Eben Li, Jianqiang Wang, and Shaobing Xu. RINO: Accurate, Robust Radar-Inertial Odometry With Non-Iterative Estimation. *IEEE Transactions on Automation Science and Engineering*, 22: 20420–20434, 2025. 1
- [43] Haibao Yu, Yizhen Luo, Mao Shu, Yiyi Huo, Zebang Yang, Yifeng Shi, Zhenglong Guo, Hanyu Li, Xing Hu, Jirui Yuan, and Zaiqing Nie. DAIR-V2X: A large-scale dataset for vehicle-infrastructure cooperative 3D object detection. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 21329–21338, 2022. 3
- [44] Haibao Yu, Wenxian Yang, Hongzhi Ruan, Zhenwei Yang, Yingjuan Tang, Xu Gao, Xin Hao, Yifeng Shi, Yifeng Pan, Ning Sun, Juan Song, Jirui Yuan, Ping Luo, and Zaiqing Nie. V2X-seq: A large-scale sequential dataset for vehicle-infrastructure cooperative perception and forecasting. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5486–5495, 2023. 2, 3, 5

- [45] Haibao Yu, Wenxian Yang, Jiaru Zhong, Zhenwei Yang, Siqi Fan, Ping Luo, and Zaiqing Nie. End-to-End Autonomous Driving Through V2X Cooperation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(9):9598–9606, 2025. [2](#), [5](#), [6](#)
- [46] Yunshuang Yuan, Hao Cheng, Michael Ying Yang, and Monika Sester. Generating evidential BEV maps in continuous driving space. *ISPRS Journal of Photogrammetry and Remote Sensing*, 204:27–41, 2023. [2](#)
- [47] Yunshuang Yuan, Yan Xia, Daniel Cremers, and Monika Sester. SparseAlign: A Fully Sparse Framework for Cooperative Object Detection. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 22296–22305, 2025. [1](#), [2](#)
- [48] Jiaru Zhong, Haibao Yu, Tianyi Zhu, Jiahui Xu, Wenxian Yang, Zaiqing Nie, and Chao Sun. Leveraging Temporal Contexts to Enhance Vehicle-Infrastructure Cooperative Perception. In *2024 IEEE 27th International Conference on Intelligent Transportation Systems (ITSC)*, pages 915–922, 2024. [2](#), [5](#)
- [49] Jiaru Zhong, Jiahao Wang, Jiahui Xu, Xiaofan Li, Zaiqing Nie, and Haibao Yu. CoopTrack: Exploring End-to-End Learning for Efficient Cooperative Sequential Perception. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 26954–26965, 2025. [1](#), [2](#), [6](#)