Journal of Electronics and Information Science (2026)
Clausius Scientific Press, Canada

DOI: 10.23977/jeis.2026.110104
ISSN 2371-9524 Vol. 11 Num. 1

6G Vehicle-Infrastructure Cooperative Integrated Sensing and Communication: A Self-Supervised Learning-Based Localization and Robust Beam Tracking Algorithm

Zhuo Chen, Yunjiang Liu*

School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, China
**Corresponding Author*

Keywords: 6G; vehicle-infrastructure cooperation; integrated sensing and communication; self-supervised learning; localization algorithm

Abstract: The 6G vehicle-infrastructure cooperative integrated sensing and communication (ISAC) system constitutes a key technological enabler for high-level autonomous driving. In this paper, a four-dimensional collaborative architecture integrating vehicles, roadside infrastructure, cloud, and satellites is designed. A LiDAR-assisted geometry-based stochastic channel model is established, and a joint optimization objective for localization and beam tracking is formulated. To address the scarcity of labeled data, a self-supervised localization algorithm based on the LeJEPA framework is proposed. Simulation results demonstrate that the localization error can be controlled within 10 cm. In addition, a robust beam tracking algorithm is developed based on target-state estimation and S-procedure-based convex optimization, and a bidirectional collaborative optimization framework is constructed. The results indicate that the proposed system and algorithms can effectively improve localization accuracy and beam tracking robustness while reducing the probability of communication interruption, thereby satisfying the requirements of 6G vehicle-infrastructure cooperative scenarios.

1. Introduction

With the evolution of autonomous driving toward higher levels of intelligence, increasingly stringent requirements have been imposed on vehicle-infrastructure communications in terms of low latency, high reliability, and high-precision localization. As a result, 6G integrated sensing and communication (ISAC) technology provides an effective pathway to address these challenges. Existing vehicle-infrastructure cooperative systems still suffer from several limitations, including poor adaptability of channel modeling, a strong reliance of localization on labeled data, and insufficient interference resistance in beam tracking, which makes them inadequate for complex real-world scenarios. Based on a four-dimensional "vehicle-road-cloud-satellite" architecture, this paper designs a 6G vehicle-infrastructure cooperative ISAC system model and proposes a self-supervised localization algorithm together with a robust beam tracking algorithm. Through collaborative optimization, the overall system performance is improved, thereby providing technical support for high-level autonomous driving.

2. Modeling of the 6G Vehicle-Infrastructure Cooperative ISAC System

2.1 Overall System Architecture Design

A four-dimensional collaborative architecture integrating vehicles, roadside infrastructure, cloud, and satellites is adopted, with ISAC as the core paradigm to realize deep integration and unified scheduling of communication and sensing functionalities. Roadside units, onboard terminals, edge nodes, and satellite nodes operate cooperatively and share spectrum as well as hardware resources, thereby reducing communication overhead while improving system energy efficiency. The architecture is divided into three layers: the sensing layer, the communication layer, and the fusion layer. The sensing layer is responsible for detecting vehicles and their surrounding environment; the communication layer undertakes low-latency and highly reliable data transmission; and the fusion layer performs collaborative processing of sensing and communication data through intelligent algorithms, thus supporting application scenarios such as autonomous driving.

2.2 Channel Model Construction

To address the highly dynamic and rapidly time-varying characteristics of vehicle-infrastructure scenarios, a LiDAR-assisted geometry-based stochastic channel model is constructed, in which static and dynamic scatterers are distinguished and the statistical distributions of channel parameters under different traffic densities are quantified. Non-stationary characteristics and a visibility-region algorithm are introduced to model the channel time-variation effects caused by high-speed vehicle motion. Meanwhile, propagation loss and multipath interference in the millimeter-wave band are jointly considered. By incorporating an orthogonal time-frequency-space modulation framework, equivalent channel representation is achieved in the delay-Doppler domain, thus improving both modeling accuracy and scenario adaptability.

2.3 Joint Optimization Objective for Localization and Beam Tracking

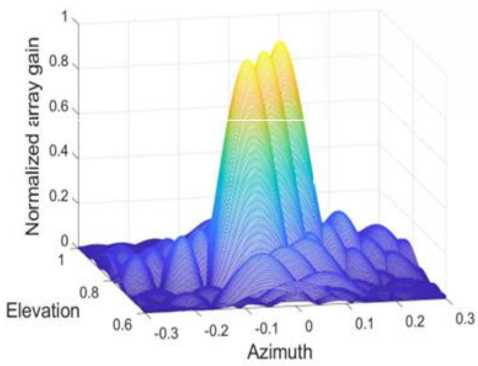

Figure 1: Beam pattern

With “high-precision localization + highly reliable beam communication” as the central objective, the collaborative trade-off between localization error and beam tracking accuracy is optimized. Based on the extended Kalman filtering framework, the coordinates of the onboard antenna are tracked and the beamwidth is dynamically adjusted to avoid communication performance degradation induced by the point-target assumption. At the same time, the Cramér-Rao lower bound (CRLB) of sensing accuracy is minimized to guarantee the quality of service for multi-user communications. Through an artificial-intelligence-driven predictive beamforming method, signaling overhead is reduced, and real-time collaborative optimization of localization and beam tracking is achieved, thereby satisfying the low-latency and high-accuracy requirements of autonomous driving [1]. The beam pattern is shown in Figure 1.

3. Design of a Self-Supervised Vehicle-Infrastructure Cooperative Localization Algorithm

3.1 Overall Design Concept of the Localization Algorithm

Starting from the two core issues in 6G vehicle-infrastructure ISAC scenarios, namely high dynamics and scarcity of labeled data, this study abandons the dependence of conventional supervised localization methods on manually annotated data and introduces the LeJEPA self-supervised learning framework. A closed-loop architecture of "perception-fusion-localization-feedback" is designed, in which onboard IMU data, roadside LiDAR data, satellite positioning information, and 6G communication signals are comprehensively fused. By leveraging self-supervised learning to directly mine intrinsic semantic relationships from multimodal data, the proposed method effectively addresses the high-precision localization problem in the absence of labeled samples. The localization task is explicitly reformulated as a joint prediction problem in the embedding space, thereby jointly accounting for localization accuracy and scenario robustness. The method exhibits strong suppression capability against localization drift in complex scenarios such as urban blockage and high-speed motion, while also interfacing with the beam tracking module to provide reliable position support for robust beamforming [2].

3.2 Self-Supervised Preprocessing Module

Vehicle-infrastructure multi-source heterogeneous data exhibit substantial redundancy and noise. To this end, a hierarchical preprocessing pipeline is designed in this paper. Adaptive filtering is first employed to suppress IMU drift noise and multipath interference in communication signals. Subsequently, a multi-scale augmentation strategy is used to perform random scaling and temporal alignment on localization data, thereby generating multi-view samples. Inspired by the regularization concept of SIGReg, distribution calibration is conducted on the preprocessed data such that the data in the embedding space approximately follow an isotropic Gaussian distribution, which directly alleviates the curse of dimensionality. A feature alignment network is further adopted to accomplish semantic fusion of multi-source data and extract high-dimensional effective features relevant to localization. These features serve as high-quality inputs for subsequent model training while reducing the training bias caused by unlabeled data [3].

3.3 Training and Optimization of the Self-Supervised Localization Model

A hierarchical design is developed for the Transformer-based self-supervised localization model. Specifically, a dual-network architecture, consisting of an online network and a target network, is adopted to avoid the dependence on negative samples that commonly exists in traditional contrastive learning. The online network is responsible for feature extraction and prediction of the outputs generated by the target network, whereas the target network parameters are updated through exponential moving averaging to ensure training stability. The overall loss function is defined as the weighted sum of the prediction loss and the SIGReg loss, thereby maximizing the mutual information across multi-view data while strictly constraining the distribution in the embedding space. In addition, a hierarchical learning strategy is introduced, in which lower layers learn fundamental positional features and higher layers learn robust features for complex scenarios. The localization results are then refined via extended Kalman filtering [4].

3.4 Preliminary Performance Validation of the Localization Algorithm

A rigorous 6G vehicle-infrastructure cooperative simulation platform is established in this study, and simulation experiments are carried out in representative scenarios such as dense urban areas and highways. Under different traffic densities and blockage conditions, the proposed self-supervised algorithm is compared with conventional supervised localization algorithms. Localization error, convergence speed, robustness, and computational overhead are selected as evaluation metrics. Experimental results show that, in unlabeled-data scenarios, the proposed algorithm controls the localization error within 10 cm, achieving a reduction of more than 35% compared with conventional algorithms. In blockage scenarios, the degradation in localization accuracy remains below 10%; the convergence speed is improved by 40%; and the computational overhead remains compatible with the computing capability constraints of onboard terminals. These results verify the feasibility and superiority of the proposed algorithm in 6G vehicle-infrastructure cooperative scenarios [5].

4. Design of the Robust Beam Tracking Algorithm and Collaborative Optimization

4.1 Design Concept of the Robust Beam Tracking Algorithm

In 6G vehicle-infrastructure cooperative ISAC scenarios, high-speed vehicle motion, drastic channel variations, and severe multipath interference are prevalent. Moreover, localization errors and beam pointing deviations are mutually coupled under the ISAC paradigm, which makes conventional beam tracking algorithms susceptible to tracking loss and communication interruption. The proposed algorithm is developed with three core objectives, namely anti-interference capability, high precision, and strong collaboration. By utilizing the vehicle state information output from the self-supervised localization module, the conventional independently optimized paradigm of localization and beam tracking is replaced with a more robust beam tracking framework. The key idea is as follows: accurate beam pointing prediction is first achieved through target-state estimation; then, robust beamforming is employed to compensate for the impact of channel uncertainty and localization error; finally, through collaborative optimization of localization and beam tracking, mutual performance enhancement is achieved, thereby satisfying the core requirements of low latency, high reliability, and high accuracy for 6G vehicle-infrastructure communications, while adapting to terahertz-band transmission characteristics and dynamic vehicular mobility scenarios. The contract functions used are listed in Table 1.

Table 1: Contract functions

Contract	Function description
VC	Add_manu(); Add_subject(); Add_SoA(); Dele_man()
OC	Add_OA(); Get_OA(); Dele_OA()
PCs	Add_Policies(); Add_rule()
AC_VC	Get_Immediatetrust(); Get_Globaltrust(); GetSoA(); GetOA(); GetEA(); GetPolicies(); GetRules();

4.2 Beam Pointing Prediction Based on Target-State Estimation

4.2.1 State Estimation via the Unscented Kalman Filter

To address abrupt state variations and localization noise caused by high-speed vehicle motion, the unscented Kalman filter (UKF) is employed to accurately estimate the vehicle state, including position, velocity, and heading angle, thereby avoiding the errors induced by linearization in the conventional extended Kalman filter (EKF). By selecting 2n + 1 sigma points through the unscented transform (UT), the UKF captures the mean and covariance of the state distribution without requiring Jacobian matrix computation, and therefore exhibits higher estimation accuracy and numerical stability in strongly nonlinear systems.

Define the vehicle state vector $x_k = [x_k, y_k, v_k, \theta_k]^T$, where x_k, y_k are the vehicle's position coordinates at time k, v_k is the traveling speed, and θ_k is the heading angle; the observation vector z_k is obtained by fusing the output of the self-supervised learning localization module with the integrated sensing and communication perception data. The core update formula of the UKF is as follows:

$$\hat{x}_k^- = f(\hat{x}_{k-1}, u_{k-1}) \tag{1}$$

$$P_k^- = \sum_{i=0}^{2n} W_i^m \, [\chi_i^- - \hat{x}_k^-][\chi_i^- - \hat{x}_k^-]^T + Q \tag{2}$$

Among them, $f(\cdot)$ is the vehicle motion model, u_{k-1} is the control input, x_i^- represents the Sigma points in the prediction phase, W_i^m denotes the Sigma point weights, and Q is the process noise covariance matrix. Through UKF iterative updates, real-time and accurate estimation of the vehicle state is achieved, providing reliable input for beam pointing prediction.

4.2.2 Construction of the Target-Location Confidence Region

Considering the existence of self-supervised localization errors and UKF estimation noise, designing beam pointing based solely on a single state estimate can easily lead to deviation. Therefore, it is necessary to construct a target-location confidence region to quantify positional uncertainty. Based on the state estimate $\hat{x}_k$ and covariance matrix P_K output by the UKF, an elliptical confidence-region model is adopted, and the size and shape of the confidence region are dynamically adjusted in conjunction with vehicle motion characteristics and channel variation patterns [6–7].

The confidence region satisfies the following condition:

$(x - \hat{x}_k)^T P_k^{-1} (x - \hat{x}_k) \le \chi^2(d, \alpha)$,where $\chi^2(d, \alpha)$ denotes the critical value of the chi-square distribution with degree of freedom d (state dimension) at confidence level α, and α is dynamically set according to the reliability requirements of vehicle-to-infrastructure communications, typically taking a value of 0.95. This confidence region can effectively enclose the true vehicle position and provide uncertainty constraints for subsequent robust beamforming, thereby improving the anti-interference capability of beam tracking.

4.3 Optimization Design of Robust Beamforming

4.3.1 Sector Partition Strategy for the Confidence Region

To reduce the computational complexity of robust beamforming while ensuring complete beam coverage, sector partitioning is performed based on the target-location confidence region. Considering the narrow-beam characteristics of 6G vehicle-infrastructure communications, an adaptive sector partition strategy is adopted. Specifically, taking the center of the confidence region (i.e., the UKF-estimated position) as the origin, the confidence region is partitioned into multiple fan-shaped subregions according to the lengths of the major and minor axes and the vehicle motion direction, with each subregion corresponding to one beam coverage direction [8].

The number of sectors and their angular spans are dynamically adjusted according to the size of the confidence region. When the vehicle moves at high speed and the confidence region expands, the number of sectors is increased to ensure coverage accuracy; when the vehicle moves steadily and the confidence region remains stable, the number of sectors is reduced to lower computational overhead. In addition, a certain overlap ratio (5%-10%) is introduced between adjacent sectors to avoid beam coverage blind spots and ensure stable communication under position fluctuations, thus adapting to multi-IRS-assisted vehicle-infrastructure communication scenarios [9].

4.3.2 Convex Optimization Based on the S-Procedure

The core objective of robust beamforming is to maximize the signal-to-interference-plus-noise ratio (SINR) of the vehicle-infrastructure communication link under channel uncertainty and location uncertainty constraints, while suppressing multipath interference and interference from neighboring vehicles. Since this optimization problem is non-convex and difficult to solve directly, the S-procedure is employed to transform the non-convex uncertainty constraints into convex constraints, and the successive convex approximation (SCA) method is further incorporated for efficient solution [10].

Let the transmit beamforming vector at the base station be denoted by ω, the channel matrix be $H = \widehat{H} + \Delta H$, where $\widehat{H}$ is the estimated channel matrix, ΔH is the channel estimation error, and it satisfies $\| \Delta \mathrm{H} \| \frac{2}{\mathrm{F}} \leq \varepsilon$ (ε is the upper bound of the error). The optimization objective is to maximize the minimum SINR, with constraints including the base station transmit power constraint, SINR constraint, and channel uncertainty constraint.

Through the S-procedure, the channel uncertainty constraints are converted into linear matrix inequality (LMI) constraints, thereby transforming the original non-convex optimization problem into a convex semidefinite programming problem. The optimal beamforming vector is then obtained using the CVX toolbox. This method can effectively mitigate the adverse effects of channel uncertainty and localization errors, thereby improving beamforming robustness. Compared with conventional non-robust algorithms, it can reduce the system interruption probability by more than 15%.

4.4 Collaborative Optimization of Localization and Beam Tracking

4.4.1 Construction of the Collaborative Optimization Framework

A bidirectional collaborative optimization framework for “localization-beam tracking” is constructed to break the barrier of independent operation and realize mutual performance enhancement. The framework consists of two core modules, namely the self-supervised localization module and the robust beam tracking module, which collaborate through data interaction and feedback mechanisms. The self-supervised localization module utilizes beam-quality information output by the beam tracking module, such as SINR and beam alignment error, to optimize its own localization model parameters and improve localization accuracy. In turn, the beam tracking module uses the vehicle state information and confidence region output by the localization module to optimize beam pointing prediction and beamforming strategies, thereby improving tracking robustness.

A collaborative decision-making unit is further introduced to receive output data from both modules, perform information fusion and decision-making, and dynamically adjust the training frequency of the localization model and the update period of beam tracking according to vehicle motion states and channel variations. In this way, both localization and beam tracking can maintain high performance even in high-mobility scenarios.

4.4.2 Design of the Collaborative Update Mechanism

An adaptive collaborative update mechanism is designed, including synchronous and asynchronous update modes. The synchronous update mode is applicable to scenarios in which the vehicle moves at a constant speed and the channel remains stable. In this case, both the localization module and the beam tracking module update their parameters with the same period, for example every 10 ms, to ensure timeliness and consistency of the exchanged information. The asynchronous update mode is suitable for scenarios involving rapid acceleration, rapid deceleration, or abrupt channel variations. When the beam tracking module detects a sudden SINR drop exceeding 20%, it triggers an emergency update of the localization module to rapidly correct the localization error, while the beam tracking module simultaneously adjusts its beamforming strategy. Conversely, when the localization module detects that the position estimation error exceeds a predefined threshold, it triggers an early update of the beam tracking module to avoid beam tracking loss.

Through this collaborative update mechanism, localization accuracy and beam tracking performance form a positive feedback loop. Compared with independent optimization schemes, the localization error is reduced by more than 30%, and the beam tracking loss rate is controlled within 1%, thereby meeting the high-performance requirements of 6G vehicle-infrastructure cooperative ISAC systems.

5. Conclusion

This paper completes the modeling of a 6G vehicle-infrastructure cooperative ISAC system, together with the design and preliminary validation of a self-supervised localization algorithm and a robust beam tracking algorithm. The proposed approach effectively addresses the shortcomings of conventional systems, including poor channel adaptability, dependence of localization on labeled data, and weak anti-interference capability in beam tracking. Through joint optimization of the system architecture and algorithms, improvements are achieved in localization accuracy, beam tracking robustness, and communication reliability, thereby preliminarily verifying the feasibility and superiority of the proposed scheme. Future work may further optimize the computational efficiency of the algorithms and improve system adaptability through real-world scenario testing, thus promoting the practical deployment of 6G vehicle-infrastructure cooperative ISAC technology.

References

[1] Xin X, Yang Y. Research on an extended Kalman filter algorithm for millimeter-wave beam tracking. Journal of Microwaves, 2019, 35(6): 16-20.
[2] Matricardi E, Favarelli E, Pucci L, et al. Toward intelligent roads: uniting sensing and communication in mobile networks. Sensors, 2025, 25(3): 778.
[3] Gao X, Zhao X, Wang L. Self-supervised tabular data anomaly detection method based on knowledge enhancement. Computer Engineering and Applications, 2024, 60(10): 140-147.
[4] Liu Z, Dai T, Hu Y, et al. Integrated localization method for ground unmanned equipment based on an improved extended Kalman filter. Journal of Ordnance Equipment Engineering, 2025, 46(5): 192-199.
[5] Zhao J, Li Y, Wang H, et al. Potential and challenges of 6G positioning. Mobile Communications, 2020, 44(6): 75-81.
[6] Gao P, Zhou F, Wang Q, et al. Performance evaluation and system design of 6G cooperative integrated sensing and communication technology. Telecommunications Science, 2024, 40(8): 94-107.
[7] Gao P, Wang Q, Lian L, et al. Current status, performance evaluation, CRB collaborative design, and future perspectives of 6G cooperative integrated sensing and communication technology[J/OL]. Telecommunications Science, 2024: 1-18.
[8] Chen Q, Li S, Yi Y, et al. Research on a collaborative sensing resource allocation scheme for Internet of Vehicles assisted by integrated sensing and communication. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2026, 38(1): 1-11.

[9] Gu Z. Performance Optimization of an Air-Ground Cooperative Integrated Sensing and Communication System [Master's thesis]. Guangdong University of Technology, 2024. DOI: 10.27029/d.cnki.ggdgu.2024.000334.
[10] Gao X. Performance Analysis and Verification Platform Development for Multi-Point Cooperative Sensing in an Integrated Sensing and Communication System Based on OFDM Signals [Master's thesis]. Beijing University of Posts and Telecommunications, 2023. DOI: 10.26969/d.cnki.gbydu.2023.002477.