Optimizing Perception Capabilities of Autonomous Vehicles Through V2I Late Fusion Using Kalman Filter

Pranav Ramesh Bidare, MS Robotics and Autonomous Systems (EE)
Mentor: Dr. Junfeng Zhao
School of Electrical, Computer and Energy Engineering

Introduction

This work presents a Kalman Filter-based Vehicle-to-Infrastructure (V2I) object fusion framework designed to enhance perception in autonomous vehicles (AVs), particularly in urban environments prone to occlusion. While onboard LiDAR systems offer high-precision depth sensing, they can struggle to detect adversarial vehicles obstructed by nearby traffic or infrastructure. To overcome this, we leverage a roadside monocular camera running the CAROM algorithm and fuse its output with onboard LiDAR data using a dual-update Kalman Filter. This fusion strategy enables the ego vehicle to detect objects at an earlier time and a greater distance away, improving reaction time and safety in complex and dynamic scenarios.

Motivation

- Enhancing Safety in Autonomous Driving: Urban environments frequently involve occlusions, making it difficult for AVs to reliably detect all road users.
- Overcoming Sensor Limitations: Onboard LiDAR may fail to detect adversarial vehicles due to obstructions.
- Maximizing the Value of V2I Infrastructure: RSU Units may provide a complementary and unobstructed viewpoint.
- Advancing Late Fusion Techniques for AVs: Fuse LiDAR and RSU camera data using a robust late fusion strategy.

Scenario

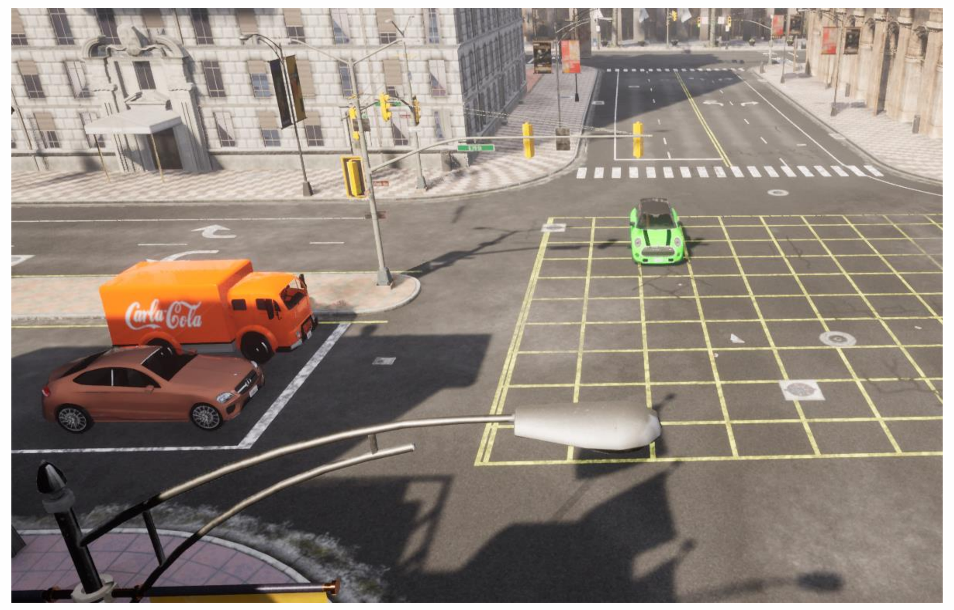

Fig 1: Snapshot of the scenario

- The scenario is taken from an actual incident reported by Waymo, where the Waymo Car failed to detect an adversary vehicle (due to occlusion) running the red light at an intersection.
- In our simulation, the ego vehicle waits at a red light while an adversarial vehicle runs through the intersection—often missed by onboard sensors due to occlusion.

Methodology

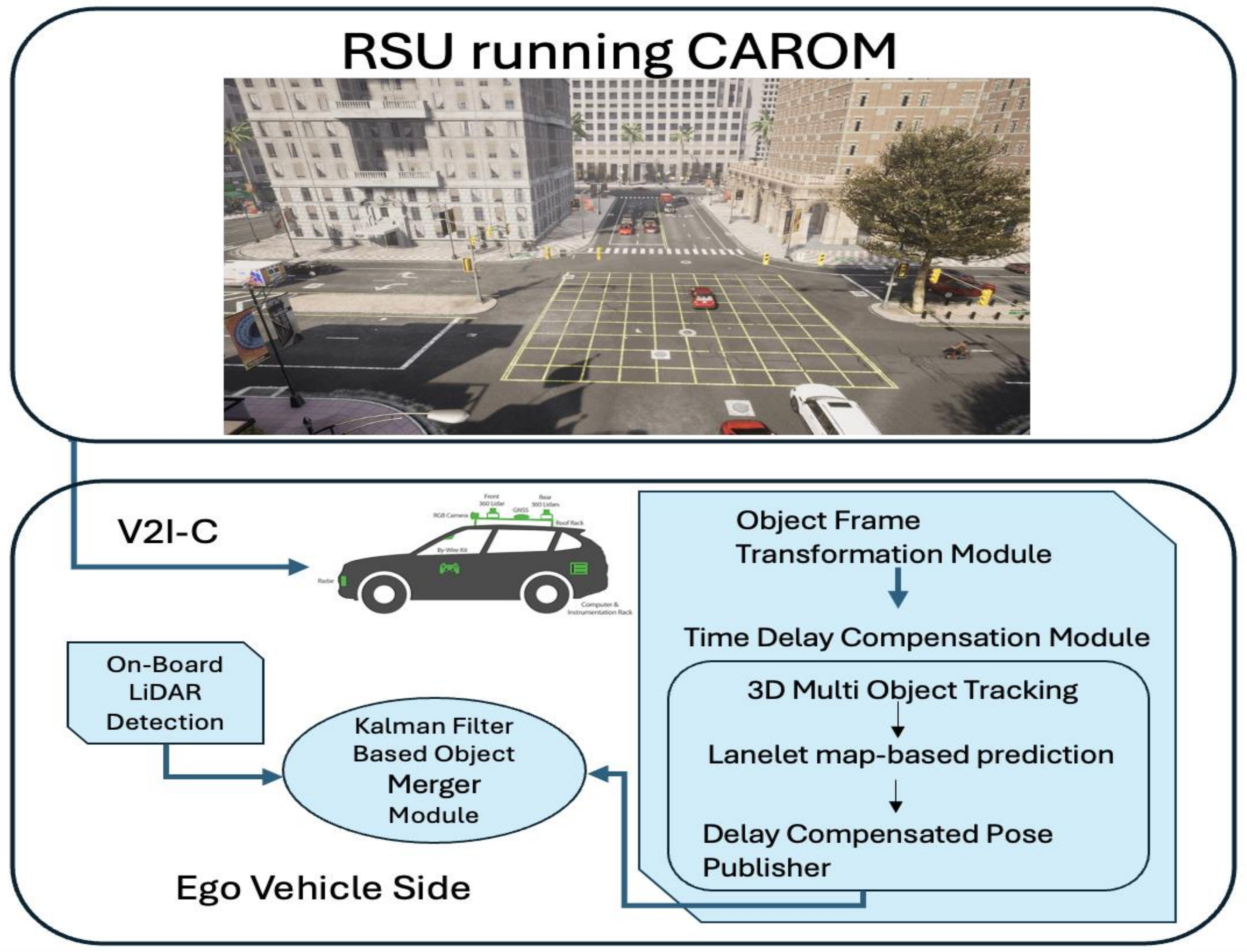

Fig 2: Block diagram of the proposed framework

- Our system fuses asynchronous sensor data from a roadside monocular camera and onboard LiDAR using a dual-update Kalman Filter.
- CAROM [1], the RSU-based 3D object detector, operates at 5 FPS and introduces 80–120ms of processing delay.
- CenterPoint, the LiDAR-based detector, provides high-accuracy object detection at 10 FPS.
- To align both data streams, RSU detections are transformed into the ego vehicle's base link frame using ROS TF-based coordinate transformation. A time delay compensation module predicts future poses of RSU-detected objects using multi-object tracking and lanelet-based trajectory forecasting.
- The Kalman Filter employs a bicycle motion model to estimate the object's state, defined as x=[x, y, z, vx, vy, psi], and performing two updates per cycle—first with LiDAR measurements to refine accuracy, followed by an update using the camera data for improved robustness. A final prediction step estimates the object's next state for real-time fusion.
- The entire pipeline is implemented in CARLA Simulator, where we evaluate it in the described scenario.

Results

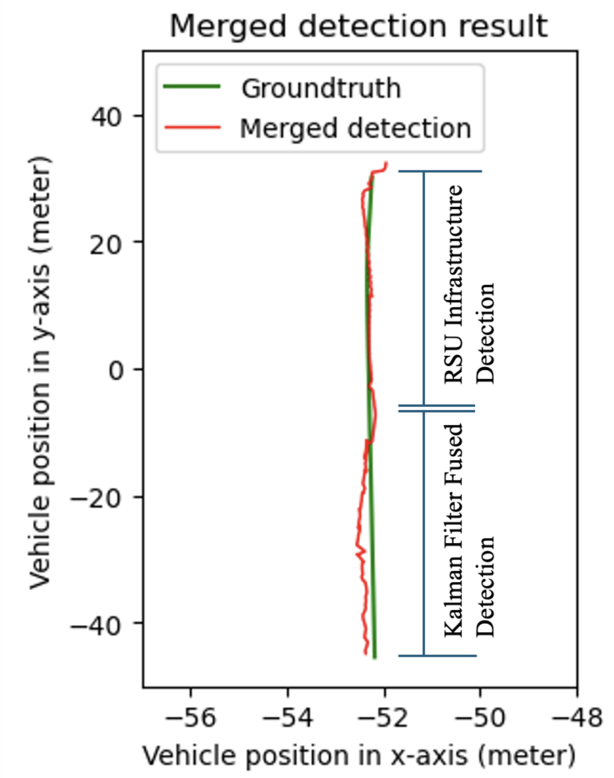

Fig 3: KF Fused detection result

Method	IoU	RMSE
Baseline [2]	0.410	2.48m
LiDAR Only KF	0.449	2.58m
Camera Only KF	0.377	2.70m
LiDAR + Camera KF	0.498	2.02m

Table 1: Results of our proposed framework

The baseline IoU-based merger achieves an IoU of **0.410** and RMSE of **2.48m**. Our proposed fusion framework achieves a superior IoU of **0.498** and RMSE of **2.02m**. Using the proposed V2I fusion framework, the ego vehicle was able to detect the adversary vehicle from **76m** away compared to just **21.6m** when relying solely on onboard perception without infrastructure assistance.

Conclusion

Our findings demonstrate that strategically leveraging infrastructure sensor data, in conjunction with predictive modeling, can substantially improve the perception capabilities, situational awareness, and overall safety of autonomous vehicles in complex environments.

Abbreviations

CAROM: CARs On the Map monocular vehicle localization algorithm
RSU: Road-Side Unit
KF: Kalman Filter
IoU: Intersection over Union
RMSE: Root Mean Square error

Reference

[1] Carom vehicle localization ad traffic scene reconstruction from monocular cameras on road infrastructures, IEEE ICRA 2021
[2] Enhanced Cooperative Perception Through Asynchronous V2I Framework with Delay Mitigation for Connected and Automated Vehicles, arXiv 2025

MORE
Masters Opportunity for Research in Engineering

ASU Ira A. Fulton Schools of Engineering
Arizona State University