



Cooperative LiDAR-Aided Self-localization of CAVs in Real Urban Scenario

Akif Adas^(✉), Luca Barbieri, Pietro Morri, Simone Mentasti, Satyesh Awasthi,
Stefano Arrigoni, Edoardo Sabbioni, and Monica Nicoli

Politecnico di Milano, Milan, Italy
akif.adas@polimi.it

Abstract. In the rapidly advancing realm of Connected Autonomous Vehicles (CAVs), achieving reliable and precise positioning is of paramount importance. This paper presents a comprehensive approach integrating vehicular sensing, communication, and advanced filtering techniques to enhance vehicle positioning in urban areas. By leveraging LiDAR point clouds along with a light and accurate object detector, we create cohesive environmental sensing that improves situational awareness in autonomous systems. Central to our methodology is the integration of the Labeled Multi-Bernoulli Mixture (LMBM) filter, which offers a probabilistic framework for dynamic state estimation in environments characterized by high uncertainty and clutter. In turn, enhanced object locations are exploited as anchors for vehicular self-localization via an Extended Kalman filter (EKF). Our experimental results show that the proposed solution significantly enhances vehicular positioning accuracy.

Keywords: Cooperative sensing · Kalman filtering · Bernoulli tracking

1 Introduction

Global Navigation Satellite Systems (GNSSs) have widespread application in Connected Autonomous Vehicle (CAV) scenarios. However, conventional GNSS systems suffer from multipath propagation in urban environments, leading to high positioning errors [8] that can not tolerated in safety-related vehicular services. A promising solution to overcome this problem is represented by collaborative sensing systems that leverage the presence of multiple actors in the same driving area to enhance positioning accuracy thanks to the coherent fusion of localization measurements acquired by nearby vehicles, as done in Implicit Cooperative Positioning (ICP) [2].

This work is financed by the European Union—NextGenerationEU project MOST (National Sustainable Mobility Center), grant # CN00000023, Italian Ministry of University and Research Decree n. 1033—17/06/2022, Spoke 9.

© The Author(s) 2024
G. Mastinu et al. (Eds.): AVEC 2024, LNME, pp. 510–516, 2024.
https://doi.org/10.1007/978-3-031-70392-8_73

This paper introduces a novel method for cooperative LiDAR-based vehicular self-positioning that leverages data sharing among vehicles to enhance localization accuracy. The innovative aspect of this method lies in its ability to address and mitigate key challenges such as false detections, object-measurement association, and measurement-measurement association which must be handled for real-world implementations. By integrating LiDAR data from various vehicles, the method ensures a more reliable and accurate vehicle positioning. Validation was conducted using a real-world dataset where multiple vehicles approached a zebra crossing on a two-way urban road featuring stationary road elements at the crossing. The results demonstrate the method's effectiveness in improving positional accuracy and robustness in complex urban environments, underscoring its potential for practical applications in autonomous driving and advanced driver-assistance systems.

The rest of the paper is organized as follows. In Sect. 2, the cooperative sensing-based self-positioning solution is described. Section 3 presents the details of the experiment and provides a quantitative assessment, followed by concluding remarks in Sect. 4.

2 Cooperative Self-positioning

The proposed method fundamentals on implicit cooperative sensing and positioning. Vehicles equipped with LiDAR sensors generate point clouds. These point clouds encompass useful information for autonomous driving such as the location of road elements. In this end-to-end method, the object detector proposed in Sect. 2.1 clusters separable 3D points in the body reference system and identifies them as objects. Among these detected objects, there exist noise-derived detections. Besides, the other two onboard units are actively engaged, namely GNSS and Inertial Measurement Unit (IMU). Here, the GNSS positioning is enhanced with a Satellite-Based Augmentation System (SBAS), if available, which requires Line-Of-Sight (LOS) conditions with at least two geosatelites. Otherwise, Single-Point Positioning (SPP) is considered.

Additionally, vehicular heading is required for LiDAR-GNSS calibration. Some approaches used for the calculation of vehicular heading are visual odometry, dual-antenna GNSS localization, and fusion of GNSS and IMU [4]. Simply, we used single-antenna GNSS fixes to calculate vehicle heading and then calibrated the rotational component of LiDAR and GNSS reference frames.

While the proposed method can be implemented in a centralized or decentralized way, this paper focuses on centralized cooperation. In turn, the following data is exchanged with Road Side Unit (RSU) via Vehicle-to-Infrastructure (V2I) communications: detected object locations in the local framework and GNSS position fix. RSU synchronizes the aggregated data and projects objects onto the global reference system. At this step, a Multi-Target Tracking (MTT) filter tracks the static road elements in the presence of clutter. Here, vehicles are considered different sensors making observations in the same area and a special form of Labeled Multi-Bernoulli Mixture (LMBM), also known as multi-sensor

LMBM, is utilized. This operation returns object locations with lower uncertainty to be considered as anchors for the subsequent steps. It is worth noting that, the accuracy and precision of target locations in the global reference frame are tightly correlated to the number of vehicles as the higher the number of vehicles, the better the accuracy. These objects are back-propagated to each vehicle via V2I communications. Within this transmission, GNSS position fix, detected targets with clutters, and augmented object locations are available at each vehicle. Subsequently, vehicles individually employ a custom Extended Kalman Filter (EKF) given in Sect. 2.3 and leverage their global positioning.

2.1 Grid-Based Spatial Detector

The Grid-Based Spatial Detector (GSD) employed in this work is designed to isolate sets of points in the 3D space that objects have generated. The process is initiated by inserting a point cluster at a known position within the point cloud to mitigate segmentation noise. Subsequently, the point cloud is projected onto a 2D space and segmented to generate a density map. This density map is subsequently filtered to reduce noise-induced segments. At the core of this object detector is a density-based clustering method, namely density-based spatial clustering of application with noise (DBSCAN) [3] which is utilized to generate clusters in the density map. In turn, the inserted point cloud has been recognized among all detections and the segmentation bias has been eliminated by shifting clusters. Consequently, clusters with small dimensions are averaged and represented as single-point objects, ensuring accurate and reliable object detection.

2.2 Cooperative LMBM

MTT is a critical process performed in autonomous systems, for detecting and tracking multiple objects in a dynamic environment. An advanced MTT method is LMBM [5], efficiently handling measurement-target data association and clutter. This filter relies on a multi-Bernoulli birth model that initiates Bernoulli components at predefined locations.

The projection of detected objects on a global reference system depicts a multi-sensor configuration with independent imaging sensors. Likewise, the detected objects are independent. Therefore, Bernoulli components are independent and follow separate Bayesian recursions. A Bernoulli Random Finite Set (RFS) β is comprised of components at time t , $\text{RFS}_\beta \triangleq \{(T, L)_\beta, r_{e,t,\beta}, \mu_{t,\beta}, \Sigma_{t,\beta}\}$, where T_β is the birth time, L_β is a unique label assigned to Bernoulli β born at time t , and $\mu_{t,\beta}$ is the first central moment and $\Sigma_{t,\beta}$ is the covariance of Gaussian distribution. Bernoulli components follow a static motion model, namely $\mu_{t,\beta} = \mu_{t-1,\beta}$, as fixed targets in the driving scenario are considered.

For the propagation of existence probability, r_e , and marginal association probabilities over time, Gibbs's sampling [6] has been employed thanks to its capability to handle high dimensional problems. After conducting the gating,

pruning and capping steps, target states are obtained by processing the remaining Bernoulli components. These states contain the spatial attribute of objects in 2D, denoted by the object set $\mathcal{O} = \{\mathbf{x}_{o,t}\} : |\mathcal{O}| = N_o$, and their corresponding covariance to be used by the filter introduced in the following section.

2.3 Extended Kalman Filter-Based Refinement Stage

A set of vehicles $\mathcal{V} = \{1, \dots, N_V\}$ cruise in a 2D area. Each vehicle v , with $v \in \mathcal{V}$, is represented by its state $\mathbf{x}_{v,t} = [\mathbf{u}_{v,t}, \mathbf{v}_{v,t}]$ where $\mathbf{u}_{v,t}$ and $\mathbf{v}_{v,t}$ are the 2D position and 2D velocity, respectively. The vehicle's state changes over time according to a nearly constant velocity motion model [7], defined as

$$\mathbf{x}_{v,t|t-1} = \mathbf{F}\mathbf{x}_{v,t-1} + \mathbf{L}\mathbf{q}_{v,t-1}, \quad (1)$$

where $\mathbf{F} = [\mathbf{I}_2 T_s \mathbf{I}_2; \mathbf{0}_{2 \times 2} \mathbf{I}_2]$ and $\mathbf{L} = [0.5T_s^2 \mathbf{I}_2; T_s \mathbf{I}_2]$, T_s the sampling time, while $\mathbf{q}_{v,t-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_2 \sigma_{v,U}^2)$ is a zero-mean Gaussian random variable with standard deviation $\sigma_{v,U}$ modeling the acceleration uncertainty of the v -th vehicle.

EKF is employed for tracking as follows. States are predicted according to the motion model in (1) and are exchanged with RSU along with detected objects provided by GSD explained in Sect. 2.1. Subsequently, the LMBM filter (Sect. 2.2) is utilized. The leveraged object locations are then back-propagated to vehicles and object-object association between detected objects and leveraged targets is established according to the maximum likelihood criterion.

For the update step of the EKF, both GNSS and object-derived measurements are utilized to enhance vehicular positioning. The GNSS measurement of vehicle v is modelled as follows:

$$\rho_{g,t}^{(v)} = \mathbf{h}_g(\mathbf{x}_{v,t}) + \mathbf{n}_{v,t} = \mathbf{T}\mathbf{x}_{v,t} + \mathbf{n}_{v,t}, \quad (2)$$

where $\mathbf{T} = [\mathbf{I}_2, \mathbf{0}_{2 \times 2}]$, while $\mathbf{n}_{v,t} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_2 \sigma_v^2)$ is a zero-mean Gaussian random variable with standard deviation σ_v^2 characterizing the GNSS positioning accuracy of the v -th vehicle.

Any object $o \in \mathcal{O}$ with state $\mathbf{x}_{v,o,t} = [\mathbf{u}_{x,o,t}, \mathbf{u}_{y,o,t}]$ refined by the RSU is more accurate compared to the detection j , such that $\mathbf{x}_{v,j,t} = [\mathbf{u}_{x,j,t}, \mathbf{u}_{y,j,t}]$, generated by the same object on vehicle v . Leveraging this fact, so o and j , the measurement models $\rho_{r,t}^{(j,v)}$ and $\rho_{\alpha,t}^{(j,v)}$, define the distance and the clockwise angle from positive x-direction between object j and vehicle v , namely they are

$$\rho_{r,t}^{(j,v)} = \mathbf{h}_r(\mathbf{x}_{v,j,t}) + \mathbf{n}_{r,t}^{(j,v)} = \|\rho_{g,t}^{(v)} - \mathbf{x}_{v,j,t}\|^2 + \mathbf{n}_{r,t}^{(j,v)}, \quad (3)$$

$$\rho_{\alpha,t}^{(j,v)} = \mathbf{h}_\alpha(\mathbf{x}_{v,j,t}) + \mathbf{n}_{\alpha,t}^{(j,v)} = \angle(\rho_{g,t}^{(v)}, \mathbf{x}_{v,j,t}) + \mathbf{n}_{\alpha,t}^{(j,v)}, \quad (4)$$

where \angle is the angle while $\mathbf{n}_{r,t}^{(j,v)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_2 \sigma_r^2)$ and $\mathbf{n}_{\alpha,t}^{(j,v)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_2 \sigma_\alpha^2)$ are zero-mean Gaussian random variables with standard deviations σ_r and σ_α , respectively. EKF refinement exploits the measurements $\rho_{r,t}^{(j,v)}$ and $\rho_{\alpha,t}^{(j,v)}$ to improve the vehicle state. Accordingly, the state estimate of the vehicle is updated as

$$\mathbf{x}_{v,t} = \mathbf{x}_{v,t|t-1} + \mathbf{G}_t (\rho_{v,t} - \mathbf{h}(\mathbf{x}_{v,t|t-1}, \mathbf{x}_{v,1,t}, \dots, \mathbf{x}_{v,N,t})), \quad (5)$$

$$\mathbf{C}_t = \mathbf{C}_{t|t-1} - \mathbf{G}_t \mathbf{B} \mathbf{C}_{t|t-1}, \quad (6)$$

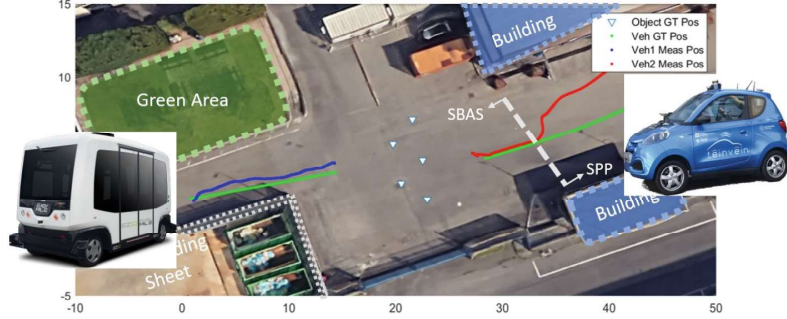


Fig. 1. Trajectories show GNSS measurements along with RTK-corrected ground truth (GT) positions, SBAS and SPP available regions along with prototypes

where $\mathbf{G}_t = \mathbf{C}_{t|t-1} \mathbf{B} (\mathbf{B} \mathbf{C}_{t|t-1} \mathbf{B}^T + \Sigma)^{-1}$ is the Kalman gain, $\mathbf{B} = [\mathbf{T}, \mathbf{J}_r, \mathbf{J}_\alpha]^T$ in which \mathbf{J}_r and \mathbf{J}_α are the Jacobian matrices extracted from (3) and (4), $\rho_{v,t} \triangleq [\rho_{g,t}^{(v)}, \rho_{r,t}^{(1,v)}, \dots, \rho_{r,t}^{(N,v)}, \rho_{\alpha,t}^{(1,v)}, \dots, \rho_{\alpha,t}^{(N,v)}]^T$, $\Sigma \triangleq \text{blkdiag}(\sigma_v^2 \mathbf{I}_2, \sigma_r^2 \mathbf{I}_N, \sigma_\alpha^2 \mathbf{I}_N)$, $\mathbf{h} = [\mathbf{h}_g(\mathbf{x}_{v,t|t-1}), \mathbf{h}_r^o, \mathbf{h}_\alpha^o]^T$ with $\mathbf{h}_\alpha^o = [\mathbf{h}_\alpha^o(\mathbf{x}_{v,1,t}), \dots, \mathbf{h}_\alpha^o(\mathbf{x}_{v,N,t})]^T$, and $\mathbf{h}_r^o = [\mathbf{h}_r^o(\mathbf{x}_{v,1,t}), \dots, \mathbf{h}_r^o(\mathbf{x}_{v,N,t})]^T$ which are range and angle measurements calculated using augmented objects.

3 Experimentation and Results

To assess the performance of the proposed method, we collected real-world data through several experiments using two different instrumented prototype vehicles, an electric light quadricycle and a prototype bus as shown in Fig. 1, equipped with comprehensive sensor systems. Collected data includes RTK-corrected and uncorrected GNSS measurements as well as LiDAR point clouds. Five objects detectable by LiDAR were used to enhance positioning, aiming to evaluate localization improvement compared to ICP. The setup, shown in Fig. 1, featured both vehicles converging towards an area with the deployed objects. For further elaboration on the experimental campaign and dataset we refer to [1].

During the implementation of the proposed method, an assumption of lossless and zero-latency V2I communication has been made. Data collected by sensors were aggregated and processed at a central computational unit offline. As concerns the computational cost, the bottleneck is LMBM in which the number of measurement-target association hypotheses increases exorbitantly. While this problem can be overcome by investigating lighter MTT algorithms, we limited the number of hypotheses and Gibbs's samples and initiated Bernoullis in the center of the surveillance area, so near to objects.

Figure 2 shows the Cumulative Density Function (CDF) of the vehicle positioning error considering ICP with [1] and without multipath compensation [2], raw GNSS ego vehicle position, proposed method, referred to as Cooperative Self-Positioning (CSP), along with Circular Error Probability with 95% confidence (CEP95). When CEP95 is concerned, CSP has a positioning improvement

of 68% and 66% compared to GNSS for veh1 and veh2 respectively. Besides, CSP has similar performance to the benchmark ICP which assumes perfect association among detected objects, making it unemployable in real applications.

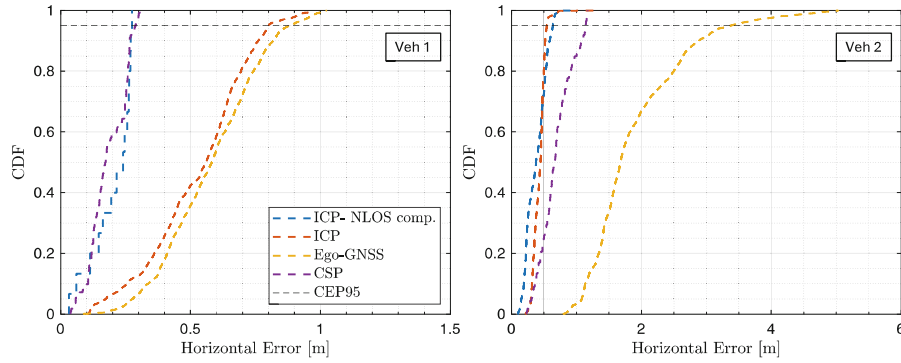


Fig. 2. Performance comparison of CSP against ICP with and without NLOS compensation and Ego-GNSS.

4 Conclusion

This paper proposed a cooperative sensing-aided positioning method relying on environmental perception with LiDAR sensors. This model-based algorithm handles the uncertainty of a dynamic environment by exploiting collaborative sensing and employing the LMBM filter. This method has been assessed in a real-world application and its effectiveness has been reported. The planned future work is to establish V2I communications and integrate RSU as both an observer and the central computational unit.

References

1. Adas, A., et al.: LiDAR-aided cooperative localization and environmental perception for CAVs. In: 2024 27th International Conference on Information Fusion (FUSION) (2024, to appear)
2. Barbieri, L., Tedeschini, B.C., Brambilla, M., Nicoli, M.: Deep learning-based cooperative lidar sensing for improved vehicle positioning. *IEEE Trans. Signal Process.* **72**, 1666–1682 (2024)
3. Hahsler, M., Piekenbrock, M., Doran, D.: dbscan: fast density-based clustering with R. *J. Stat. Softw.* **91**(1), 1–30 (2019)
4. Juang, J.C., Lin, C.F.: A sensor fusion scheme for the estimation of vehicular speed and heading angle. *IEEE Trans. Veh. Technol.* **64**(7), 2773–2782 (2015)
5. Reuter, S., Vo, B.T., Vo, B.N., Dietmayer, K.: The labeled multi-Bernoulli filter. *IEEE Trans. Signal Process.* **62**(12), 3246–3260 (2014)

6. Robert, C.P., Casella, G.: The two-stage Gibbs sampler. In: Robert, C.P., Casella, G. (eds.) Monte Carlo Statistical Methods, pp. 337–370. Springer, New York (2004). https://doi.org/10.1007/978-1-4757-4145-2_9
7. Thrun, S., Montemerlo, M.: The graph slam algorithm with applications to large-scale mapping of urban structures. *Int. J. Robot. Res.* **25**(5–6), 403–429 (2006)
8. Zhu, N., Marais, J., Bétaille, D., Berbineau, M.: GNSS position integrity in urban environments: a review of literature. *IEEE Trans. Intell. Transp. Syst.* **19**(9), 2762–2778 (2018)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

