Sparsity-Aware Occupancy Grid Mapping for Automotive Driving Using Radar-LiDAR Fusion

Zhai, Peiyuan; Joseph, Geethu; Myers, Nitin Jonathan; Önen, Çağan; Pandharipande, Ashish

DOI
10.1109/SENSORS60989.2024.10785054

Publication date
2024

Document Version
Final published version

Published in
2024 IEEE Sensors, SENSORS 2024 - Conference Proceedings

Sparsity-Aware Occupancy Grid Mapping for Automotive Driving Using Radar-LiDAR Fusion

Peiyuan Zhai
EEMCS, TU Delft
The Netherlands
p.zhai@tudelft.nl

Geethu Joseph
EEMCS, TU Delft
The Netherlands
g.joseph@tudelft.nl

Nitin Jonathan Myers
ME, TU Delft
The Netherlands
n.j.myers@tudelft.nl

Çağan Önen
NXP Semiconductors
The Netherlands
cagan.onen_1@nxp.com

Ashish Pandharipande
NXP Semiconductors
The Netherlands
ashish.pandharipande@nxp.com

***Abstract*—We tackle the problem of estimating a binary occupancy grid map by fusing point cloud data from LiDAR and radar sensors for automotive driving perception. To this end, we introduce two sparsity measurement models for fusion, formulating occupancy mapping as a sparse binary vector reconstruction problem. The first model jointly estimates a common map from all measurements, while the second assumes a shared map and an innovation component for each modality's measurements. We use the pattern-coupled sparse Bayesian learning algorithm to recover maps, leveraging the inherent sparsity and spatial dependencies in automotive occupancy maps. Numerical experiments on the RADIATE public dataset show that our fusion-based approach improves mapping accuracy compared to single-modality and high-level fusion mapping algorithms.**

***Index Terms*—Sparse Bayesian learning, feature-level fusion, multi-modal sensing, RADIATE dataset.**

2024 IEEE SENSORS | 979-8-3503-6351-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/SENSORS60989.2024.10785054

I. Introduction

In assisted and autonomous driving, occupancy grid mapping (OGM) is used to generate a representation of the vehicle's surroundings, typically using LiDAR or radar measurements. These maps are essential for subsequent driving tasks, such as object tracking and path planning [1], [2]. We address the sensor fusion problem for OGM, focusing on combining point cloud data from LiDAR and radar to improve the detection of targets and identify drivable areas.

The state-of-the-art occupancy mapping algorithms focus on either single sensing modalities or high-level fusion of maps generated by different sensing modalities. A common single-modality occupancy mapping algorithm is the inverse sensor model, where the occupancy probability of each cell is updated recursively based on measurements. The algorithm has also been extended to multi-modality sensor cases and fused in the feature or decision level [3], [4]. However, the model often struggles with conflicting measurements and does not adequately explore the spatial relationships between occupancy grids. Another popular approach for OGM is the kernel-based methods such as Gaussian Process OGM and its variants [5]–[7], for which the decision-level heterogeneous sensor fusion can be performed probabilistically [8]. The kernel-based methods exploit the spatial dependencies but cannot account for the underlying sparsity of the occupancy status.

In typical automotive applications, the OGM is sparse as the obstacles constitute small continuous areas of the entire environment and are measured only at their boundaries. The sparsity-aware OGM model is introduced in [9], where a block sparse occupancy map is estimated using the pattern-coupled sparse Bayesian learning (PCSBL) algorithm [10]. However, the work in [9] is limited to single-modality sensors, either LiDAR or radar, and does not address issues of data unreliability or calibration errors.

In this paper, we study the LiDAR and radar point cloud fusion problem using a sparse OGM model. To effectively utilize information from diverse sensors, we use feature-level fusion rather than high-level fusion [11]. Our main contributions are twofold. We present two feature-stage LiDAR-radar sensor fusion models for OGM, namely common sparse and common-innovation sparse (CIS) fusion models. Then, using the real-world RADIATE public dataset [12], we demonstrate that our fusion models achieve superior drivable area detection and comparable target detection performance relative to the single-modality sparse OGM [9] and high-level fusion of the two single-modality sparse maps. Moreover, our CIS model can dynamically adjust to account for unreliable sensor data while fusing, thereby improving the map accuracy.

II. Sensor fusion for sparsity-aware mapping

This section introduces the sensor signal models, the mapping problem, and presents our fusion algorithms.

A. Signal model and problem formulation

The LiDAR and radar data are first preprocessed by removing points outside the mapping area [9]. Let M_L and M_R be the remaining number of LiDAR and radar points. To define the map, we divide the area of interest into N grid cells. Then, the OGM problem is to estimate a vectorized occupancy map $\mathbf{x} \in \mathbb{R}^N$, where $x[i]$ denotes the probability that the i^{th} grid cell is occupied. The vector $\mathbf{x}$ is sparse, i.e., contains few zeros, in a typical automotive setup.

To formulate the relationship between $\mathbf{x}$ and sensor measurements, we identify the occupied and free cells associated with each measurement. For LiDAR measurements, the cells along the line connecting the ego vehicle are free, and the cell closest to the reflection point is occupied [13]. As radar has a lower angle resolution than a LiDAR, a beam is used instead of a line to identify free and occupied cells [9]. Each point cloud yields two linear measurements: free cell occupancies sum

to a low value (near zero), while occupied cell occupancies sum to a high value [9]. This leads to a linear model for LiDAR measurements $\mathbf{y}_L \in \mathbb{R}^{2M_L}$ and radar measurements $\mathbf{y}_L \in \mathbb{R}^{2M_R}$, denoted by

$$\mathbf{y}_L = \mathbf{A}_L\mathbf{x}_L + \mathbf{n}_L \quad \text{and} \quad \mathbf{y}_R = \mathbf{A}_R\mathbf{x}_R + \mathbf{n}_R. \tag{1}$$

Here, $\mathbf{x}_L, \mathbf{x}_R \in \mathbb{R}^N$ are the single-modality maps, and $\mathbf{n}_L$ and $\mathbf{n}_R$ are the measurement noise. Our goal is to combine $\mathbf{y}_L$ and $\mathbf{y}_R$ to estimate the underlying occupancy map $\mathbf{x}$.

B. Common sparse fusion

Our first method uses a common sparse map-based model, which assumes that $\mathbf{x}_L = \mathbf{x}_R = \mathbf{x}$. This leads to a new model that concatenates the LiDAR and radar measurements as

$$\begin{bmatrix}\mathbf{y}_L \\ \mathbf{y}_R\end{bmatrix} = \begin{bmatrix}\mathbf{A}_L \\ \mathbf{A}_R\end{bmatrix}\mathbf{x} + \begin{bmatrix}\mathbf{n}_L \\ \mathbf{n}_R\end{bmatrix}. \tag{2}$$

We assume different noise levels for the LiDAR and radar measurements, $\mathbf{n}_L \sim \mathcal{N}(0, \sigma_L^2\mathbf{I})$ and $\mathbf{n}_R \sim \mathcal{N}(0, \sigma_R^2\mathbf{I})$, to account for differences between the measurement modalities.

To solve for a sparse OGM $\mathbf{x}$ from (2) with unknown σ_L^2 and σ_R^2, we turn to a Bayesian algorithm, called PCSBL [9], [10]. It imposes a hypothetical two-layer hierarchical prior on the map. The first layer is a Gaussian prior with hyperparameter $\boldsymbol{\alpha}$,

$$p(\mathbf{x}|\boldsymbol{\alpha}) = \prod_{n=1}^{N} \mathcal{N}(x_n|0, \delta_n^{-1}), \tag{3}$$

Here, the precision $\delta_n = \alpha_n + \beta \sum_{m\in\mathcal{N}_n} \alpha_m$ of the nth cell depends on its hyperparameter and that of the direct neighbors denoted by $\mathcal{N}_n$. In addition, β is an algorithm parameter. The second layer is a Gamma hyperprior model with parameters a, b,

$$p(\boldsymbol{\alpha}) = \prod_{n=1}^{N} \text{Gamma}(\alpha_n|a, b), \tag{4}$$

where a larger hyperparameter a value leads to a sparser map and vice versa. Using the above prior model, we estimate $\mathbf{x}$ from (2) using the expectation-maximization (EM) algorithm [9]. The EM algorithm iteratively estimates the hyperparameters $\boldsymbol{\alpha}$ and noise variances σ_L^2, σ_R^2. For more details on the algorithm, we refer the reader to [10].

By jointly using the measurements, the algorithm effectively resolves conflicts due to inconsistent measurements, leading to a better map than single modality maps.

C. Common-Innovation Sparse (CIS) support fusion

In the second model, we softly fuse LiDAR and radar measurements based on the sensor's reliability. To accommodate any possible errors in the single-modality maps, we assume that $\mathbf{x}_L = \mathbf{x} + \mathbf{x}_{\Delta L}$ and $\mathbf{x}_R = \mathbf{x} + \mathbf{x}_{\Delta R}$. Here, $\mathbf{x}_{\Delta L}$ and $\mathbf{x}_{\Delta R}$ are the error collectors that capture false positives. We substitute $\mathbf{x}_L$ and $\mathbf{x}_R$ in (1) to write

$$\begin{bmatrix}\mathbf{y}_L \\ \mathbf{y}_R\end{bmatrix} = \begin{bmatrix}\mathbf{A}_L & \mathbf{A}_L & \mathbf{0} \\ \mathbf{A}_R & \mathbf{0} & \mathbf{A}_R\end{bmatrix} \begin{bmatrix}\mathbf{x} \\ \mathbf{x}_{\Delta L} \\ \mathbf{x}_{\Delta R}\end{bmatrix} + \begin{bmatrix}\mathbf{n}_L \\ \mathbf{n}_R\end{bmatrix}. \tag{5}$$

To compute the sparse map $\mathbf{x}$ from (5), we resort to EM-based sparse Bayesian learning with a different hierarchical prior. Here, $\mathbf{x}$ follows (3), and $\mathbf{x}_{\Delta L}$, $\mathbf{x}_{\Delta R}$ follows a Gaussian prior with precision $\alpha_{\Delta L}$, $\alpha_{\Delta R}$, respectively [14]. The entries of α, $\alpha_{\Delta L}$, and $\alpha_{\Delta R}$ follow Gamma prior in (4) controlled by the hyperparameter a, taking values a_C, a_L, and a_R, respectively. If both sensors are reliable, we choose high a_L and a_R values, enforcing $\mathbf{x}_{\Delta L} = \mathbf{x}_{\Delta R} = \mathbf{0}$. The resulting algorithm is similar to the common sparse fusion algorithm. If LiDAR (or radar) is known to be less reliable, the corresponding hyperparameter a_L (or a_R) is set to smaller than a_R (or a_L) to allow the error to be captured by $\mathbf{x}_{\Delta L}$ (or $\mathbf{x}_{\Delta R}$). This approach does not discard less reliable sensors, but instead selectively incorporates their measurements by adjusting the sparsity of the error collectors, outperforming single-modality maps.

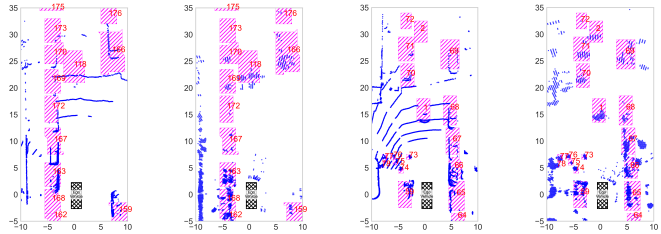
III. Numerical Results

We use two error metrics, Intersection over Bounding Box (IoBB) and Angular Scan Normalized Mean-Squared Error (AS-NMSE) [9], to quantify the performance of the algorithms. IoBB quantifies the target detection via the ratio of the detected cells to the occupied cells in the ground truth. AS-NMSE evaluates the NMSE in lengths of unoccupied areas across multiple directions, indicating the drivable area detection.

Our results are shown in two sample scenes from the RADIATE dataset [12]: City-3-0 frame-275 and City-3-7 frame-120. The baseline algorithms are single-modality LiDAR OGM [9], single-modality radar OGM [9], and logical OR operator-based fusion of the two single-modality maps. The hyperparameters are set to $a = 0.5$ and $b = 10^{-4}$. For CIS fusion OGM, when both sensors are reliable, we choose $a_C = 0.5$ and $a_L = a_R = 1.3$. If LiDAR (or radar) is less reliable, we change $a_L = 0.53$ (or $a_R = 0.53$).

(a) City-3-0 frame-275 (b) City-3-7 frame-120

Fig. 1. The view from the ego-vehicle in the RADIATE dataset scenes.



(a) SC1: LiDAR (b) SC1: Radar (c) SC2: LiDAR (d) SC2: Radar

Fig. 2. Point clouds (blue points) of scenes: (a)-(b) City-3-0 frame-275 and (c)-(d) City-3-7 frame-120. The purple-shaded boxes are the ground truth.

TABLE I
RESULTS OF SCENE CITY-3-0 FRAME 275

Method	IoBB													Detected targets	AS-NMSE
	118	159	162	163	166	167	168	169	170	172	173	175	176		
LiDAR OGM	0.088	0.233	0.000	0.241	0.038	0.185	0.042	0.000	0.000	0.040	0.000	0.000	0.000	7 / 13	0.052
Radar OGM	0.418	0.209	0.765	0.185	0.397	0.148	0.125	0.204	0.250	0.040	0.000	0.000	0.139	11 / 13	0.211
OR fusion OGM	0.429	0.395	0.765	0.426	0.405	0.296	0.167	0.204	0.250	0.080	0.000	0.000	0.139	11 / 13	0.198
Common sparse fusion OGM	0.396	0.349	0.765	0.204	0.359	0.241	0.125	0.102	0.192	0.060	0.019	0.000	0.139	12 / 13	0.038
CIS fusion OGM	0.396	0.349	0.765	0.204	0.359	0.241	0.125	0.102	0.192	0.060	0.019	0.000	0.139	12 / 13	0.039

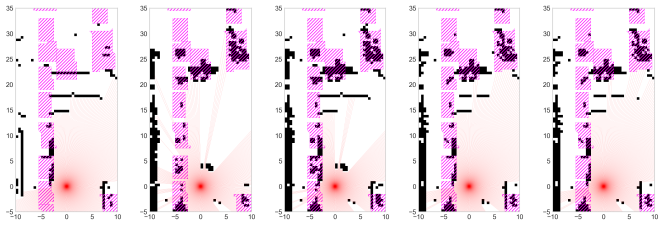
TABLE II
RESULTS OF SCENE CITY-3-7 FRAME 120

Method	IoBB																		Detected targets	AS-NMSE
	1	2	59	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78		
LiDAR OGM	0.200	0.000	0.080	0.067	0.258	0.256	0.200	0.109	0.000	0.133	0.074	0.000	0.500	0.500	0.250	0.250	0.500	1.000	15 / 18	0.130
Radar OGM	0.333	0.475	0.020	0.133	0.303	0.077	0.380	0.127	0.377	0.378	0.389	0.167	0.000	0.000	0.000	0.750	0.500	0.000	14 / 18	0.139
OR fusion OGM	0.489	0.475	0.100	0.200	0.530	0.333	0.540	0.236	0.377	0.467	0.426	0.167	0.500	0.500	0.250	0.750	0.667	1.000	18 / 18	0.133
Common sparse fusion OGM	0.400	0.475	0.060	0.200	0.470	0.256	0.420	0.182	0.325	0.356	0.389	0.167	0.500	0.500	0.250	0.250	0.667	1.000	18 / 18	0.123
CIS fusion OGM	0.267	0.475	0.080	0.133	0.379	0.077	0.360	0.200	0.364	0.378	0.370	0.167	0.500	0.500	0.250	0.250	0.667	1.000	18 / 18	0.083

A. Scene City-3-0 frame-275

This scene represents a regular city road with thirteen vehicles either parked or moving on the road, as shown in Fig. 1 (a). The LiDAR and radar point clouds are shown in Fig. 2 (a) and (b), respectively. The mapping results are in Fig. 3, and IoBB and AS-NMSE are listed in Table I.

The results show that our common sparse and CIS fusion OGM algorithms have the best AS-NMSE value compared to the baseline algorithms. While our methods match the IoBB values of the OR fusion, they also eliminate false positives in front of the ego-vehicle via the feature-level LiDAR-radar fusion. This approach leads to a lower AS-NMSE value, indicating superior drivable area detection performance. The low false alarm rate may have led to the detection of target 173 using its LiDAR point cloud measurements with a small IoBB value. In contrast, this target was missed by both single-modality LiDAR OGM and OR fusion OGM. Consequently, our common sparse and CIS fusion approaches detect the most targets (12 out of 13). Additionally, our methods show similar performances because the CIS fusion algorithm uses large hyperparameter values ($a_L > 1$ and $a_R > 1$) to make $\mathbf{x}_{\Delta L}$ and $\mathbf{x}_{\Delta R}$ close to zero, thereby making (2) and (5) equivalent.



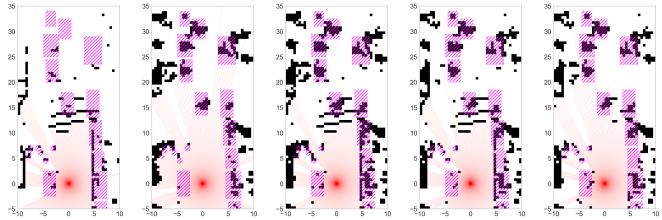
(a) LiDAR (b) Radar (c) OR fusion (d) Common sparse fusion (e) CIS fusion

Fig. 3. Occupancy grid maps of City-3-0 frame-275. The black cells are the detected targets. The red lines represent the drivable area lines.

B. Scene City-3-7 frame-120

This scene shows a crossroad with ten parked vehicles, two moving vehicles (Targets 1-2), and six pedestrians (Targets 73-78) as shown in Fig. 1 (b). The LiDAR and radar point clouds are shown in Fig. 2 (c) and (d). For this scene, the LiDAR measurements in the dataset show a small calibration error and significant ground-scatter error, making it less reliable. Exploiting the side information that LiDAR is less reliable, we set the LiDAR hyperparameter $a_L = 0.53 < a_R$. The resulting maps and error metrics are shown in Fig. 4 and Table II.

Here, all fusion methods detect more targets than the single-sensor maps and successfully identify all the targets. The performance difference of the fusion algorithms is in the AS-NMSE value. Compared to the OR fusion, the common sparse fusion reduces a few false positives, slightly improving the AS-NMSE value. However, the CIS fusion map largely reduces the LiDAR ground scatters while maintaining target detection, leading to the best AS-NMSE. These results highlight the potential of the CIS fusion model to achieve better performance than the OR and common sparse fusion models when one of the sensors is less reliable. Since the CIS model estimates more unknowns, i.e., the map $\mathbf{x}$ and the error collectors $\mathbf{x}_{\Delta L}$ and $\mathbf{x}_{\Delta R}$, it has a higher runtime than other algorithms.



(a) LiDAR (b) Radar (c) OR fusion (d) Common sparse fusion (e) CIS fusion

Fig. 4. Occupancy grid maps of City-3-7 frame-120. The black cells are the detected targets. The red lines represent the drivable area lines.

IV. CONCLUSION

We presented two feature-level sensor fusion models for sparse occupancy grid mapping: the common sparse and CIS fusion models. Results with scenes from the RADIATE dataset showed that when the sensors are reliable, both models enhance target detection compared to single-sensor approaches and improve drivable area detection relative to the baseline methods. When one sensor is less reliable, CIS surpasses other methods, albeit at a higher computational cost. A promising direction for future research includes joint camera-LiDAR-radar fusion, exploring the sparse structure of the map.

REFERENCES

[1] S. Steyer, C. Lenk, D. Kellner, G. Tanzmeister, and D. Wollherr, "Grid-based object tracking with nonlinear dynamic state and shape estimation," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 7, pp. 2874–2893, Jul. 2020.

[2] A. Pandharipande, C.-H. Cheng, J. Dauwels, S. Z. Gurbuz, J. Ibanez-Guzman, G. Li, A. Piazzoni, P. Wang, and A. Santra, "Sensing and machine learning for automotive perception: A review," *IEEE Sens. J.*, vol. 23, no. 11, pp. 11 097–11 115, Mar. 2023.

[3] A. Elfes and L. Matthies, "Sensor integration for robot navigation: Combining sonar and stereo range data in a grid-based representataion," in *Proc. IEEE Conf. Decision Control*, vol. 26, Dec. 1987, pp. 1802–1807.

[4] J. Sock, J. Kim, J. Min, and K. Kwak, "Probabilistic traversability map generation using 3D-LIDAR and camera," in *Proc. IEEE Int. Conf. Robot. and Autom.*, May 2016, pp. 5631–5637.

[5] S. T. O'Callaghan and F. T. Ramos, "Gaussian process occupancy maps," *Int. J. Robot. Res.*, vol. 31, no. 1, pp. 42–62, Jan. 2012.

[6] J. Wang and B. Englot, "Fast, accurate Gaussian process occupancy maps via test-data octrees and nested Bayesian fusion," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2016, pp. 1003–1010.

[7] K. Doherty, J. Wang, and B. Englot, "Bayesian generalized kernel inference for occupancy map prediction," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2017, pp. 3118–3124.

[8] M. P. Gerardo-Castro, T. Peynot, and F. Ramos, *Laser-Radar Data Fusion with Gaussian Process Implicit Surfaces*. Springer International Publishing, 2015, pp. 289–302.

[9] Ç. Önen, A. Pandharipande, G. Joseph, and N. J. Myers, "Occupancy grid mapping for automotive driving exploiting clustered sparsity," *IEEE Sens. J.*, 2024.

[10] J. Fang, Y. Shen, H. Li, and P. Wang, "Pattern-coupled sparse Bayesian learning for recovery of block-sparse signals," *IEEE Trans. on Signal Process.*, vol. 63, no. 2, pp. 360–372, 2015.

[11] D. L. Hall and J. Llinas, Eds., *Handbook of multisensor data fusion*, ser. The Electrical engineering and applied signal processing. Boca Raton, FL: CRC Press, 2001.

[12] M. Sheeny, E. De Pellegrin, S. Mukherjee, A. Ahrabian, S. Wang, and A. Wallace, "Radiate: A radar dataset for automotive perception in bad weather," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2021, pp. 1–7.

[13] J. E. Bresenham, "Algorithm for computer control of a digital plotter," *IBM Syst. J.*, vol. 4, no. 1, pp. 25–30, Jul. 1965.

[14] M. E. Tipping, "Sparse Bayesian Learning and the Relevance Vector Machine," *J. Mach. Learn. Res.*, vol. 1, no. 3, pp. 211–244, 2001.