FUTURITY PUBLISHING

Section of the conference: Computer science, computing and automation

Sekcja konferencji: Informatyka, obliczenia i automatyzacja

How to cite: Bondariev, I. (2025). Spatial Mapping of Motion Detection Zones Through the Fusion of Radar and Optical Data. *World Conference on Emerging Science, Innovation and Policy 2025*. Futurity Research Publishing. https://doi.org/10.5281/zenodo.20644196

Spatial Mapping of Motion Detection Zones Through the Fusion of Radar and Optical Data

Illia Bondariev

1 Master of Arts in International Relations, Technical Product Manager, Amazon.com Services LLC, Sunnyvale, USA, https://orcid.org/0009-0003-9768-9480

Accepted: Jul 9, 2025 **|** **Published:** July 20, 2025 **|** **Language:** English

Abstract: The objective of the study was to substantiate a model for the spatial mapping of motion detection zones based on the fusion of radar and optical data. The methodology encompassed bibliographic synthesis, logical modeling, sensor synchronization, projection of the data onto a common coordinate grid, and computation of a cell reliability index. The results are presented as a model that partitions space into stable, conditionally reliable, uncertain, and blind zones; in the worked example, the share of reliable coverage was 65%. The proposed approach can be applied in traffic monitoring, perimeter security, industrial safety, and the preliminary auditing of sensor system configurations.

Keywords: radar-camera fusion, motion detection, spatial mapping, occupancy grid, sensor reliability.

Introduction

Spatial mapping of motion detection zones is an important applied task in computer vision and sensor-based monitoring, one that requires not only the recording of individual moving objects but also the determination of which areas of a monitored territory remain under stable sensor coverage. For transportation hubs, warehouse logistics, security systems, and robotic navigation, such maps make it possible to distinguish reliably monitored zones from areas where motion may go undetected because of lighting conditions, field-of-view occlusion, radar shadow, or geometric constraints on observation.

Optical sensors provide a highly detailed representation of a scene and allow the contours and semantic characteristics of objects to be identified; however, their effectiveness depends substantially on illumination, weather conditions, the presence of shadows, and perspective distortion. Radar sensors, in turn, directly provide information on range, azimuth, and Doppler velocity, and they remain operational in fog, rain, and darkness. At the same time, radar data are characterized by lower density and limited potential for semantic interpretation. For this reason, current research treats radar and cameras as complementary rather than interchangeable sources of information (Yao et al., 2024; Wang et al., 2025).

The relevance of the topic stems from the shift from local object detection toward the construction of integrated spatial representations of the environment, particularly in the form of a bird's-eye view or occupancy grids. Such approaches make it possible to combine radar points with optical features within a single coordinate space, to assess the reliability of individual areas, and to generate a spatial confidence layer that is interpretable for an operator (Baek et al., 2024; Sun et al., 2024). However, most existing studies focus predominantly on three-dimensional object detection, whereas the mapping of motion detection zones requires a distinct methodological approach and specialized models of spatial assessment.

Research Aim and Research Tasks

The aim of the study is to substantiate a compact model for the spatial mapping of motion detection zones based on the fusion of radar and optical data within a common coordinate grid.

To achieve this aim, the following research tasks were defined: 1) to synthesize current approaches to radar-optical fusion and determine their suitability for zonal mapping; 2) to describe the coordinate alignment of radar measurements and optical features within a common plane; 3) to propose a formula for an integral reliability index R for each cell; 4) to substantiate a scale for partitioning space into stable, conditionally reliable, uncertain, and blind zones; 5) to demonstrate the model on a computational scenario and outline indicators for subsequent field verification.

The methodological foundation of the study comprised a theoretical and applied approach, a bibliographic synthesis of recent work on radar-optical fusion, logical modeling of sensor interaction, a coordinate representation of the monitored territory, and computational modeling of the zonal map. In the first stage, the directions of data fusion were systematized and the requirements for a common representation plane were defined. In the second stage, a cell-based grid model was constructed and the contributions of the radar and optical channels were described. In the third stage, the reliability index R and the scale for interpreting cells were specified. In the fourth stage, a computational demonstration was carried out for a 60 × 40 m area. The work is presented not as a completed field experiment but as a methodological basis for subsequent empirical validation.

Research Results

The theoretical basis of the study is the concept of multilevel sensor fusion, according to which data of differing physical natures are transformed into a common spatial representation before a decision is made about the presence of motion. Within this logic, the radar channel is regarded as a source of metric information on range, azimuth, and Doppler velocity, whereas the optical channel provides contour, texture, and semantic detail of the scene. It is precisely the combination of these functions that justifies the transition from the detection of individual objects to the assessment of the quality of sensor coverage across the entire monitored territory.

A bibliographic review of current sources made it possible to identify five leading directions in the development of this problem. The first direction concerns the classification of early, feature-level, intermediate, and late fusion, which determines the processing stage at which signals from different sensors are combined (Yao et al., 2024; Qian et al., 2025). The second direction addresses the transition to a bird's-eye view representation, in which radar and optical features are aligned within a common plane and can be compared at the level of zones rather than only at the level of object bounding boxes (Cai et al., 2024; Fent et al., 2025). The third direction encompasses the construction of dynamic occupancy grids, in which the cell becomes the minimal unit for assessing space, motion, and semantic state (Jang et al., 2024). The fourth direction relates to tracking and temporal stability, since a single sensor activation does not always confirm sustained motion within a zone (Baek et al., 2024; Cheng et al., 2024). The fifth direction focuses on enhancing the robustness of the radar channel and on adapting the fusion weights to conditions of visibility, noise, and point sparsity (Sun et al., 2024; Wang et al., 2025).

For the task of spatial mapping, it is significant that the directions noted above do not yield a ready-made map of motion zones but are oriented primarily toward 3D detection, classification, or object tracking. For this reason, the present study proposes an intermediate applied model in which the fusion results are interpreted as a cell reliability index rather than as a final object class. This approach makes it possible to show the operator not only the location of probable motion but also the quality of sensor confidence in each part of the monitored space.

On the basis of the theoretical review, three levels of radar and optical data fusion were distinguished. The first level involves the matching of completed detections, in which radar points or tracks are confirmed by optical bounding boxes. The second level operates on features, which makes it possible to combine radar velocity and spatial depth with visual contours. The third level forms a common map in the form of a grid, in which each cell receives its own motion estimate. For spatial mapping, the third level is the most suitable, since it shifts the focus from the individual object to the quality of territorial coverage and directly corresponds to the first task of the study.

Table 1

Functional contribution of the radar and optical channels to the mapping of motion zones

Component	Radar data	Optical data	Role in the zone map
Geometry	range, azimuth	image coordinates, contours	projection onto a single grid
Motion	Doppler velocity	inter-frame displacement	confirmation of motion
Semantics	limited data on object class	visual class and shape	interpretation of detections
Robustness	better performance in fog, rain, and darkness	better performance under clear visibility	compensation for sensor limitations
Mapping output	sparse but metric signal	dense but visibility-dependent signal	reliability index at the cell level

Source: developed by the author

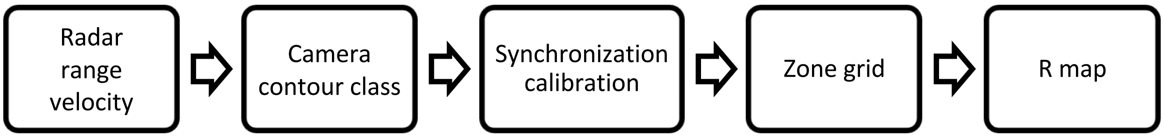
As shown in Table 1, the radar channel performs a metric and dynamic function, whereas the optical channel refines the contour and content of the scene. This means that fusion should not be reduced to the mechanical summation of activations. It is more appropriate to apply a cell-based model in which each sensor contributes evidence to a particular area of space.

The monitored territory is represented as a set of cells g_i. For each cell at time t, the radar motion probability $P_r(g_i,t)$, the optical motion probability $P_o(g_i,t)$, and the temporal stability coefficient $S(g_i,t)$ are computed. The integral reliability index is defined by the formula $R(g_i,t) = w_r \cdot P_r(g_i,t) + w_o \cdot P_o(g_i,t) + w_s \cdot S(g_i,t)$, where $w_r + w_o + w_s = 1$. In the baseline scenario, the values $w_r = 0.40$, $w_o = 0.40$, and $w_s = 0.20$ were adopted. Under conditions of poor visibility, the weight of the radar channel is increased, whereas under good visibility, the weight of the optical channel is increased.

Cells with $R \geq 0.80$ are interpreted as stable zones; cells with $0.60 \leq R < 0.80$, as conditionally reliable; cells with $0.40 \leq R < 0.60$, as uncertain; and cells with $R < 0.40$, as blind or problematic. Such a scale does not replace the final decision of the security system, but it creates an explanatory layer: the operator can see where sensor confidence is high, where it is formed by only one channel, and where additional sensor coverage is required. This accomplishes the task of formalizing the zonal interpretation of space.

Figure 1

Pipeline of radar-optical fusion for spatial motion-zone mapping



Source: author's own development

The structural model encompasses five stages [Figure 1]. First, radar and optical data are synchronized by timestamp. Next, spatial calibration is performed, in which radar coordinates and optical features are converted into a common plane. Following this, motion evidence is accumulated for each cell, the index R is computed, and a zone map is generated. This logic is consistent with current approaches in which data from different sensors are projected into a common spatial representation to enhance perceptual robustness (Xiao et al., 2023; Cai et al., 2024). In this way, the model sequentially addresses the second and third tasks of the study.

For the computational example, a monitored area of 60 × 40 m was taken with a grid step of 2 m. The total number of cells is 30 × 20 = 600. If the radar effectively covers 480 cells, the camera covers 520, and the intersection of the two channels comprises 420 cells, then the baseline share of joint coverage equals 420/600 × 100 = 70.0%. If, after fusion, 390 cells receive $R \geq 0.60$, the operational share of reliable coverage equals 390/600 × 100 = 65.0%. The remaining 210 cells, or 35.0% of the area, require either an adjustment of the sensor mounting angle, a correction of activation thresholds, or the installation of an additional sensor node.

The practical value of the model lies in its capacity for the preliminary auditing of sensor configurations. If the map reveals a large number of blind cells, the mounting height or angle of the camera, the radar sector, or the grid density can be adjusted accordingly. If redundant coverage is detected, the system can be optimized with respect to computational load. The zonal map is therefore not only a means of presenting results but also a design tool.

For the subsequent empirical validation of the model, it is appropriate to record the share of stable, conditionally reliable, uncertain, and blind cells; the accuracy of motion confirmation by radar and camera; the share of false activations; the temporal stability of the index R; and the effect of illumination, precipitation, and occlusions on changes in the weighting coefficients. Such indicators will make it possible to verify not only the fact of object detection but also the reliability of the spatial coverage of the monitored territory.

Conclusions

The study synthesizes current approaches to radar-optical fusion and demonstrates the appropriateness of the transition from object-level detection to cell-based mapping of the monitored territory. The radar channel is identified as a source of metric and dynamic features, and the optical channel as a source of contour and semantic detail; together, they support the construction of a common coordinate grid.

An integral reliability index R is proposed, combining the radar motion probability, the optical motion probability, and the temporal stability of the cell. On this basis, a partition of space into stable, conditionally reliable, uncertain, and blind zones is substantiated. The computational scenario for a 60 × 40 m area demonstrated that, with joint coverage of 70.0% of cells, the operational share of reliable coverage after fusion amounts to 65.0%.

The results obtained can be applied to the auditing of sensor coverage, the selection of sensor installation positions, and the explanation of incomplete motion detection. Further research should be directed toward the field validation of the model, the refinement of weighting coefficients for varying weather conditions, and the comparison of multiple spatial grid scales.

References

Baek, S., Kim, J., & Yi, K. (2024). Robust tracking and detection based on radar camera fusion filtering in urban autonomous driving. *Intelligent Service Robotics*, *17*, 1125–1141. https://doi.org/10.1007/s11370-024-00563-0

Cai, G., Chen, F., & Guo, E. (2024). IRBEVF-Q: Optimization of image–radar fusion algorithm based on bird's eye view features. *Sensors*, *24*(14), 4602. https://doi.org/10.3390/s24144602

Cheng, L., Sengupta, A., & Cao, S. (2024). Deep learning-based robust multi-object tracking via fusion of mmWave radar and camera sensors. *IEEE Transactions on Intelligent Transportation Systems*, *25*(11), 17218–17233. https://doi.org/10.1109/TITS.2024.3421339

Fent, F., Palffy, A., & Caesar, H. (2025). DPFT: Dual perspective fusion transformer for camera-radar-based object detection. *IEEE Transactions on Intelligent Vehicles*, *10*(11), 4929–4941. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10769556

Jang, H., Kim, T., Ahn, K., Jeon, S., & Kang, Y. (2024). Dynamic occupancy grid map with semantic information using deep learning-based BEVFusion method with camera and LiDAR fusion. *Sensors*, *24*(9), 2828. https://doi.org/10.3390/s24092828

Qian, H., Wang, M., Zhu, M., & Wang, H. (2025). A review of multi-sensor fusion in autonomous driving. *Sensors*, *25*(19), 6033. https://doi.org/10.3390/s25196033

Sun, X., Jiang, Y., Qin, H., Li, J., & Ji, Y. (2024). Camera-radar fusion with radar channel extension and dual-CBAM-FPN for object detection. *Sensors*, *24*(16), 5317. https://doi.org/10.3390/s24165317

Wang, H., Liu, J., Dong, H., & Shao, Z. (2025). A survey of the multi-sensor fusion object detection task in autonomous driving. *Sensors*, *25*(9), 2794. https://doi.org/10.3390/s25092794

Xiao, Y., Liu, Y., Luan, K., Cheng, Y., Chen, X., & Lu, H. (2023). Deep LiDAR-radar-visual fusion for object detection in urban environments. *Remote Sensing*, *15*(18), 4433. https://doi.org/10.3390/rs15184433

Yao, S., Guan, R., Huang, X., Li, Z., Sha, X., Yue, Y., & Yue, Y. (2024). Radar-camera fusion for object detection and semantic segmentation in autonomous driving: A comprehensive review. *IEEE Transactions on Intelligent Vehicles*, *9*(1), 2094–2128. https://doi.org/10.1109/TIV.2023.3307157