D-S EVIDENCE THEORY-DRIVEN FPGA ARCHITECTURE FOR RADAR AND VISUAL FUSION ALGORITHM

Hao Ni, Hongfei Lian, Qiuyu Liu, Hongqi Fan

College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China

Abstract: To address the issue of low real-time fusion efficiency in multi-source sensor data, this paper proposes an FPGA-based radar-vision fusion method using D-S evidence theory. By implementing a parallel architecture and pipeline optimization, the fusion latency is reduced to the microsecond level while improving resource utilization. Simulation experiments demonstrate that the proposed method offers high reliability and low latency in autonomous driving scenarios, with potential scalability to intelligent transportation systems.

Keywords: D-S evidence theory, radar-vision fusion, FPGA, autonomous driving

In autonomous driving multimodal perception systems, the fusion of millimeter-wave radar and visual sensors encounters challenges in real-time processing of heterogeneous data. While Dempster-Shafer (D-S) evidence theory enhances target detection robustness through confidence fusion, its high computational complexity impedes real-time applications. This study proposes a hardware acceleration module based on FPGA parallel architecture and pipeline optimization, achieving high-speed synthesis of dynamically conflicting evidence via the D-S algorithm. The solution delivers low-latency, highly reliable multisource perception for autonomous vehicles, addressing microsecond-level response demands in complex scenarios.

The D-S evidence theory-based radar-visual fusion algorithm is implemented in two stages. First, Basic Probability Assignment (BPA) functions are constructed for millimeter-wave radar and visual sensors respectively, tailored to their heterogeneous data characteristics. An unknown event category is incorporated to model target classification uncertainty, ensuring normalization constraints in probability allocation. Second, the D-S combination rule dynamically fuses BPA data from both modalities, with fusion weights adaptively adjusted based on quantified sensor uncertainty levels to resolve evidence conflicts. As illustrated in Figure 1, this framework enhances multi-source perception confidence and robustness through heterogeneous data modeling and probabilistic fusion mechanisms. Detailed implementation procedures are provided in Reference [2].

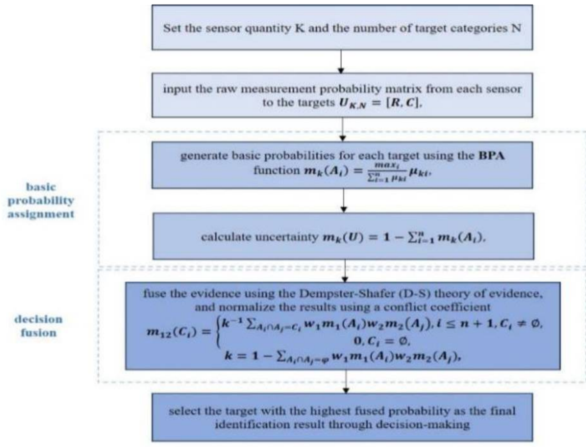

Figure 1,Algorithm Flowchart

The FPGA implementation comprises two phases: first validating the D-S algorithm logic in MATLAB, followed by hardware migration via Vitis HLS. To address real-time requirements for radar-visual fusion, a collaborative optimization strategy integrating loop unrolling and pipeline techniques is adopted. This approach establishes a parallel architecture for BPA computation across millimeter-wave radar and visual sensors, enhancing data throughput while optimizing hardware resource utilization. System validation is conducted through a three-stage process encompassing C simulation, synthesis, and co-simulation to ensure compliance with stringent latency and reliability constraints in multimodal perception scenarios.

To validate the functional correctness of the algorithm, C simulation is employed. In this case, data from relevant literature [3] is used as input. Specifically, visual sensor data originates from ImVoxelNet (WACV 2022) [4], and millimeter-wave radar data comes from PointPillars (CVPR 2019) [5], as shown in Table 1.

Table 1 - Credibility of Possible Targets for Vision Sensors and Radar Sensors（%）

Sensor	Car	Pedestrian	Cyclist	Truck
Vision Sensor	22.55	13.73	9.67	13.87
Radar Sensor	21.26	28.33	52.47	11.18

The results obtained after applying the Basic Probability Assignment (BPA) are shown in Table 2.

Table 2 - Basic Probability Assignments of Possible Targets for Vision Sensors and Radar Sensors(%)

Sensor	Car	Pedestrian	Cyclist	Truck	Uncertainty
Vision Sensor	8.50	5.18	3.65	5.23	77.45
Radar Sensor	9.85	13.13	24.31	5.18	47.53

The results after fusion are shown in Table 3.

Table 3 - Fusion Results of Vision Sensors and Radar Sensors(%)

Car	Pedestrian	Cyclist	Truck	Uncertainty	Fusion Result
13.77	14.65	23.61	7.45	40.52	Cyclist

By comparing Table 1 and Table 3, it can be observed that neither a single vision sensor nor a radar sensor can achieve a unified identification of the target based solely on its own detection results. However, by applying D-S evidence theory for fusion, the final detection result identifies the target as a cyclist, which is consistent with the original results provided in reference [2]. Furthermore, by comparing Table 2 and Table 3, it is evident that the fusion process reduces the uncertainty of the detector. Lower uncertainty indicates higher reliability of the identification.

The FPGA hardware architecture mapping was implemented through C synthesis, generating critical performance metrics. The acceleration factor of the code was calculated by comparing its execution efficiency with MATLAB simulation results, as summarized in Table 4.

Table 4 - Results of C Synthesis for Vitis HLS Code and Speedup Factor Benchmarking

Key Metrics	Target	Estimated	Uncertainty	Latency(ns)	Latency in Matlab(ms)
Synthesis Results	20.00ns	13.42ns	5.40ns	1680.00	5.301
Key Metrics	**Interval**	**BRAM**	**DSP**	**FF**	**LUT**
Synthesis Results	85	0	32	3059	5466

The analysis of Table 4 reveals that the actual clock period of 14.60 ns meets the requirement of exceeding the estimated 13.42 ns. The total fusion latency is measured at 1,680 ns, achieving a 3100× speedup factor compared to the 5.3 ms latency observed in MATLAB simulations for identical datasets. Furthermore, resource utilization metrics—including DSP slices, flip-flops (FFs), and look-up tables (LUTs)—remain significantly below the maximum thresholds of FPGA platforms typically deployed for radar-visual fusion tasks. These results validate the algorithm's capability to support high-throughput processing of large-scale inputs in autonomous driving systems while maintaining hardware efficiency.

Functional correctness and timing violations were validated through co-simulation, with the timing diagram generated during this process illustrated in Figure 2.

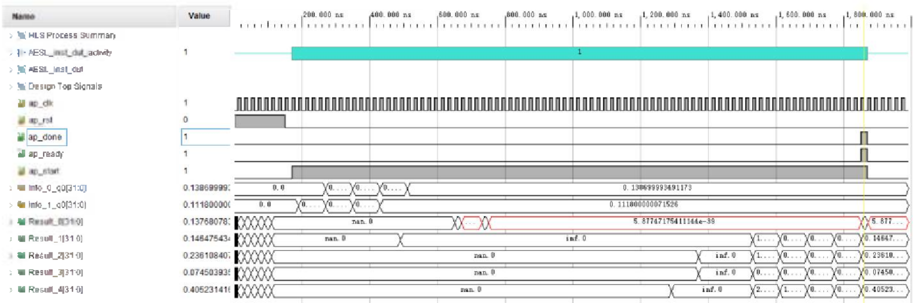

Figure 2, Timing diagram generated after co-simulation using Vitis HLS software

Figure 2 shows that the simulation results are consistent with the C simulation results, indicating that the functionality of the hardware implementation is correct. The sensor inputs are sampled accurately on the rising clock edges, and the fusion results are output within the expected cycles. Key signals (such as start, done) are stable at the clock edges, presenting no risk of metastability. This confirms the successful realization of the FPGA design implementation.

To address the challenges of low multi-sensor fusion efficiency and insufficient reliability of single-sensor environmental perception in autonomous driving, this paper proposes an FPGA-based hardware acceleration method for radar-visual fusion utilizing Dempster-Shafer (D-S) evidence theory. A hierarchical parallel pipeline architecture combined with dynamic resource scheduling strategies achieves efficient fusion of millimeter-wave radar and vision heterogeneous data, leveraging FPGA's parallel computing capabilities to compress fusion latency to microsecond levels. This solution provides all-weather reliable perception for autonomous vehicles while demonstrating extensibility to smart transportation systems and agricultural robotics domains. The proposed methodology exhibits significant industrial application prospects for multi-modal perception technologies in dynamic complex scenarios

Reference

1. Yao S., Guan R., Huang X. [et al.] Radar-Camera Fusion for Object Detection and Semantic Segmentation in Autonomous Driving: A Comprehensive Review // IEEE Transactions on Intelligent Vehicles. – 2023. – Т. 9, № 1. – С. 1-40.

2. Xie H., Pu J. Research on Information Fusion Algorithm of Multi-mode Composite Detector Based on D-S Evidence Theory // Photoelectric Technology Application. – 2017. – Т. 32, № 2. – С. 41-43.

3. Bai X., Yu Z., Zheng L. [et al.] SGDet3D: Semantics and Geometry Fusion for 3D Object Detection Using 4D Radar and Camera // IEEE Robotics and Automation Letters. – 2025. – Т. 10, № 1. – С. 828-835.

4. Rukhovich D., Vorontsova A., Konushin A. ImVoxelNet: Image to voxels projection for monocular and multi-view general-purpose 3D object detection // Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. – 2022. – С. 1265-1274.

5. Lang A.H., Vora S., Caesar H. [et al.] PointPillars: Fast encoders for object detection from point clouds // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2019. – С. 12689-12697.