Research Statement

Sai Bharadhwaj Matha
Email: bharadhwaj2299@gmail.com
Phone: +49 15510859066
LinkedIn: https://linkedin.com/in/saibharadhwajmatha
Portfolio: https://bharadhwajsaimatha.github.io/portfolio/

1. Introduction

My research bridges the fields of embedded systems and intelligent 3D perception, with a long-term goal of developing embodied agents that perceive, reason, and act autonomously in complex environments. My research began in electrical and embedded systems, where I designed power electronics and flight-embedded hardware and firmware, before evolving into advanced 3D perception, deep learning, and generative modeling. This cross-disciplinary arc has shaped my pursuit of machine intelligence grounded in both the physical and perceptual world.

Today, my focus lies in agent-centric 3D perception: how intelligent systems construct semantic, panoptic, geometric, and temporal representations of the world for reasoning and control. I aim to integrate my hardware background with cutting-edge research in spatial AI, enabling intelligent agents that are not only data-driven but also physically embedded and sensor-aware, which is critical for autonomous robotics, AR/VR, and future AGI systems, and possibly beyond.

2. Research Background

2.1 From Electromechanics to Learning Systems

My Bachelor's thesis on *Wireless Power Transfer for Solar Harvesting* involved modeling resonant coupling for energy efficiency, which foreshadowed my interest in signal interaction and dynamic systems. During my Master's degree at TH Deggendorf, I shifted my research focus to artificial intelligence, studying machine learning, deep vision architectures, and reinforcement learning. This gave me a broad foundation for 3D perception and the motivation to pursue research at the interface of hardware and intelligent perception.

2.2 Embedded Systems and Applied Robotics

As an Embedded Systems Engineer at Ideaforge (2020–22), I developed Field Oriented Control (FOC)-based ESCs, embedded firmware, PX4 autopilot integration, and GPS-denied navigation systems for UAVs. My work extended from low-level motor control to Li-ion charging hardware and real-time SLAM integration (ORB-SLAM3 with thermal/RGB-D). This hands-on experience with sensing, control, and mission-critical software taught me how tightly-coupled embedded design is with robotic autonomy.

2.3 Semantic and Panoptic 3D Perception for UAVs

At Fraunhofer IVI, I designed a novel semantic and panoptic occupancy dataset and benchmarking pipeline for advanced aerial robotics. The pipeline acts on monocular images and integrates Structure from Motion, Multi-View Stereo, 2D to 3D semantic lifting, voxelization, voxel densification, geometry-aware extrusion, mesh-based hole filling, and frustum-based culling, amassing over 50K monocular samples with semantic and panoptic ground truth 3D volumes, making it the only large-scale dense SSC and PSC dataset for UAVs. I adapted and ran ablations on SOTA models (e.g., CGFormer, Symponize3D) and am currently developing a new geometry-aware transformer architecture tailored to aerial 3D perception. Latest upgrades include the inclusion of thermal imagery and experimenting with 3D perception from thermal imagery.

2.4 Research Philosophy

Whether modeling power loss in coils or semantic affinity losses in voxels, my research consistently seeks generalizable, efficient, and explainable system representations. I believe that perception should not be a passive pipeline, but an agentic process, one that takes advantage of uncertainty, multi-modality, and learned 3D perception to interact meaningfully with the world.

3. Current and Ongoing Research

CVPR 2026 Submission: I am finalizing a novel 3D semantic-panoptic occupancy framework that combines geometry-aware voxel priors and visibility-aware semantic priors to perceive 3D complex aerial environments. The model explicitly reasons over occlusion and geometry, while the dataset is novel, dense, and approximately 6 times larger than SemanticKITTI.

MonoSpatial: I am leading a project that combines monocular RGB input with fine-tuned depth models and diffusion-based priors to estimate scene geometry and camera intrinsics. This enables agent-based spatial queries with minimal sensor assumptions.

Multi-modal Perception: I am exploring fusion architectures for 3D object detection under adverse weather (e.g., fog), training encoders on synthetic data in CARLA, and exploring early fusion methods across monocular images and raw LiDAR point clouds.

These projects showcase my growing interest in 3D perception under real-world constraints, where data scarcity, physical noise, and limited sensing require both learning and design ingenuity.

4. Future Research Vision

My goal is to unify my knowledge of hardware and perception to build intelligent, autonomous agents that understand and interact with the world in 3D. Key directions include:

- **Spatial World Models**: Multi-Modal representations that integrate semantics, geometry, and dynamics—optimized for memory and latency.

- **Embodied AI**: Architectures that couple real-time sensor data with learned priors to drive robotic behavior and spatial imagination.
- **Generative 3D Perception**: Diffusion-based and Gaussian-based models that synthesize scenes from sparse or ambiguous inputs.
- **Hardware-Aware AI**: Designing perception systems that are co-optimized with embedded hardware for low-power, on-device autonomy.

I believe AGI will not emerge from perception alone, it will require the convergence of sensors, signals, and semantics. By leveraging my embedded systems background, I aim to build AI that is not just intelligent in abstraction, but embodied in the real world.

5. Conclusion

My trajectory reflects a deep curiosity for both how systems work and how they perceive. I seek to contribute to a research environment that values both theory and application, one where embedded design and AI perception work together to produce robust, intelligent, and autonomous systems. I look forward to collaborating on projects that combine hardware, learning, and 3D spatial intelligence to advance the future of embodied AGI.

Note: References and publications available on request or at: https://bharadhwajsaimatha.github.io/portfolio/