

AVISHAI SHMARIAHU

AI Systems | Computer Vision & ML | Technical Program Leader

Los Gatos, CA | avishai_shmariahu@hotmail.com | (408) 476-4297

PROFESSIONAL SUMMARY

Senior Technical Program Manager with deep hands-on expertise across real-time Computer Vision and machine perception systems, AI/ML acceleration over custom silicon platforms for XR devices.

Working closely with architecture, algorithm, embedded software, and SoC hardware teams from early technology exploration through development and product launch, driving engineering planning and execution across highly complex multidisciplinary systems.

Strong background influencing architecture and roadmap direction while balancing product requirements, technology limitations, engineering plans, and program goals.

CORE COMPETENCIES

- **Computer Vision Architecture:** SLAM, head, hands, eye and body tracking, 3D reconstruction and Occlusions.
- Custom Silicon **Accelerators** - IP features roadmap and development program from IP to SOC to Product
- **AI/ML Inference Acceleration and optimization** - Model exporting and quantization: PyTorch, ExecuTorch, PTQ/QAT workflows
- **Embedded FW & Sensor-to-OS Pipelines** Design and Implementation
- Large scale and complexity **Program Management** involving state-of-the-art innovation
- **New Technology Introduction (NTI)** Research-to-Product Program Leadership
- **New Product Introduction (NPI)** Cross-Functional Execution in High-Ambiguity Environments
- **Power and Latency optimization**

PROFESSIONAL EXPERIENCE

Technical Program Manager, Computer Vision System

Meta Reality Labs | Sunnyvale, CA | May 2020 - Present

- **Computer Vision & AI:** Lead highly technical Computer Vision and AI/ML subsystem programs from IP Architecture to Product implementation, enabling real-time AR perception and interaction experiences across XR devices, spanning custom silicon, AI/ML Compiler, inferencing, embedded software, sensors, and machine perception systems.
- **Program Scale:** Managed 10+ parallel development plans involving 50+ engineers and architects across silicon architecture, hardware, sensors, ML acceleration, embedded software, XR algorithms, and product organizations - delivering state-of-the-art real-time, power-efficient, low-latency machine perception capabilities including 6DoF head-poses, hands, eyes, and body tracking, enabling AR experiences such as World/Chest/Head Lock, 3D reconstruction, occlusion, and hands/eyes-based inputs.
- **Applied Agentic AI:** Deep practitioner of LLMs and agentic AI workflows - hands-on with Manus and Claude Code as primary tools, piloted extensively within engineering teams.
- **AI Models at Scale:** Drove cross-functional teams to deploy optimized AI/ML inference pipelines and 20+ models across custom compute architectures, AI/ML acceleration flows, and constrained embedded platforms, enabling advanced Computer Vision experiences.
- **Technical Leadership:** Influenced and established architectural capability by working closely with architecture, algorithm, embedded software, hardware, and product teams on architecture direction, implementation strategy, technology tradeoffs, and cross-functional alignment.
- **Silicon IP Program Leadership:** Established and led the IP Readiness Program for custom silicon development across 6 IP domains (CV, GFX, Display, Compression, Audio, Security), aligning architecture, RTL design, SoC, and leadership on a shared execution framework tied to the SoC delivery schedule.
- **Planning at scale:** Established cohesive engineering OKRs, long-term roadmaps, and detailed execution plans across the NPI lifecycle — from Discovery and Concept through Build, Validation, and Launch — aligned with Semi-Annual

planning cycles, hardware revisions, and organizational goals, enabling deployment of advanced AR capabilities into product-ready platforms.

- **De-Risking:** Identified cross-functional dependencies, technology gaps, riskier paths, and execution complexities early in the development cycle, driving mitigation plans through alternate paths, architecture tradeoffs, and shift-left execution plans.
- **Optimized Implementation:** Led optimization efforts improving latency at frame-rate performance, power efficiency within budget, and reduced memory footprint while maintaining model accuracy across real-time perception systems and AI/ML acceleration flows.

Senior Technical Program Manager, DeepLens AI Devices

[AWS AI/ML \(Amazon\)](#) | East Palo Alto, CA | Sep 2018 - May 2021

- Led cross-functional programs to establish roadmap, development, and launch plans for AI/ML edge devices, cloud services, and developer console features.
- Drove execution through technical complexity, shifting priorities, and risk mitigation - maintaining 360° alignment and cohesive plans across hardware, software, and cloud teams.
- Managed AI accelerator SoC partner down-select, including NVIDIA Jetson Xavier for the AWS Panorama Appliance - a computer vision edge device running ML inference via NVIDIA TensorRT.
- Managed ODM partnerships in Asia for Panorama device hardware and embedded software co-development.

Senior TPM & Product Manager, Visual Search & Augmented Reality

[A9.com \(Amazon\)](#) | Palo Alto, CA | Nov 2016 - Aug 2018

- Led design and engineering teams to launch computer vision-powered mobile features and Deep Learning models, from problem identification through roadmap definition and production launch.
- Key launches: Scan & Pay (India offline payments), SmileCode WW launch, credit card recognition improvement (+14% accuracy), and enhanced scanner UX (+8% success rate).

Engineering Program Management Director

[SanDisk](#) | Milpitas, CA | Aug 2007 - Oct 2016

- Directed large-scale hardware/firmware development programs with strategic SoC partners, supply chain, and Asia operations, delivering SSD and flash storage products to retail and OEM markets.
- Matrix-managed 20+ program managers, optimizing NAND technology cost/performance across client SSD, MicroSD, USB, and Compact Flash product lines.

Previous Engineering & Leadership Roles

[Brocade Communications](#), [3Com](#), [Tower Semiconductor](#), [Rafael](#) | pre-2007

- Led engineering, platform, and product development programs across semiconductor, networking, embedded hardware, and defense systems.
- Managed multidisciplinary teams spanning silicon, firmware, hardware, manufacturing, and supply chain through full product lifecycle.

EDUCATION

MBA, Marketing Technion - Israel Institute of Technology, Haifa

B.Sc., Mechanical Engineering, Summa Cum Laude Technion - Israel Institute of Technology, Haifa