Zwanzigerstraße 16, Lindau, Germany
+49-15128983877
mundra.akshay15@gmail.com
amundra15.github.io
akshay-mundra

Akshay Mundra

Overview

I am a **Computer Vision Engineer** with a background in **generative AI, 3D vision & graphics**, and 5 years of work experience. I recently completed my master's degree in Visual Computing at Saarland University, where my thesis focused on a NeRF-based method for creating photorealistic 3D hand avatars from 2D images, resulting in a publication at **ICCV '23**. Currently, I am working on Diffusion Models and Visual Foundation Models for autonomous driving.

Education

Oct '20 - Aug '23 — **Saarland University** — *Saarbrücken, Germany*
M.Sc. in Visual Computing — *GPA: 1.3* [1]

Aug '14 - May '18 — **Birla Institute of Technology and Science** — *Pilani, India*
B.E. (Hons.) in Electronics and Instrumentation — *GPA: 7.91/10*

Work Experience

Sep '23 - Present — **DENSO ADAS** — *Lindau, Germany*
Computer Vision Engineer

- Delivered a proof of concept for **occupancy prediction** for autonomous driving. The transformer-based model takes multi-view RGB images as input and predicts per-voxel occupancy.
- Developing a **Vision Foundation Model** by distilling knowledge from a pre-trained **diffusion model**. The work emphasizes improving robustness to adverse weather conditions and out-of-distribution scenarios using **self-supervised learning**.

Aug '21 - Aug '23 — **Max Planck Institute for Informatics** — *Saarbrücken, Germany*
Research Assistant — Supervisor: Prof. Dr. Christian Theobalt — Thesis

- Created a **NeRF-based approach** to learn personalised 3D hand avatars from multi-view images.
- The model renders human hands in **real-time** with photorealistic details. It also models hand-pose and camera-view dependent changes in the hand texture.
- Implemented a live demo to **track** the user's hand and **render** it in real-time.
- The work has been accepted to **ICCV '23**.

Jun '18 - Sep '20 — **DreamVu Inc.** — *Hyderabad, India*
Computer Vision Engineer

- Built a Generative Adversarial Network **(GAN) based image restoration model** to remove imaging artefacts such as defocus, noise and optical specularity.
- Developed a novel structured-light based **camera calibration** method to convert optically coded images to 360° stereo panoramas, significantly enhancing the camera's imaging capabilities.
- Oversaw the camera assembly process from an imaging perspective at the company's facility in Johor Bahru, Malaysia.

Jun '17 - Dec '17 — **Computer Vision Centre** — *Barcelona, Spain*
Research Intern — Supervisor: Dr. Antonio López — Thesis

- Developed an **end-to-end perception** and navigation system for an autonomous vehicle in a virtual environment using imitation learning.
- Fine-tuned the model with real-world data to bridge the **sim-to-real domain gap**.
- Showcased the work at Smart City Expo World Congress '17.

[1] German grading system – Best: 1.0, Worst: 5.0.

Publications

[1] **Mundra, A.**, BR, M., Wang, J., Habermann, M., Theobalt, C., Elgharib, M. (2023). LiveHand: Real-time and Photorealistic Neural Hand Rendering. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Webpage Paper

Projects

Nov '20 - Jan '21 **Ray tracer in C++** Webpage

- Built a ray tracer from scratch in C++, with salient features such as acceleration structures and distribution ray tracing.
- Showcased the ray tracer in a rendering competition.

Jun '21 - Jul '21 **Automated Traffic Control Monitoring** Report

- Developed an end-to-end ML pipeline to estimate traffic density from smartphone and vehicle dashcam images.
- Pre-processed an in-the-wild dataset and trained a lightweight MobileNetV2 model on it.
- Deployed the model on an android compatible application for real-time inferencing.

Jun '21 - Jul '21 **Multi-Frame Super Resolution for Smartphone Photography** Report

- Generated bursts of low-resolution images synthetically, imitating tremors common in hand-held photography.
- Applied transfer learning to a *Residual Feature Attention* based model for multi-frame super-resolution.

Skills

Knowledge Generative AI, Diffusion Models, Deep Learning, Imaging, 3D Reconstruction, SLAM, SSL

Languages Python, C++, MATLAB, Java, C, Bash

Libraries PyTorch, TensorFlow, OpenCV, PyTorch3D, Metashape, FiftyOne, Pandas, OpenMP

Tools AWS, Jupyter, Docker, Slurm, Git, POV-Ray, Linux, CMake, LaTeX

Soft skills Problem Solving, Research, Teamwork, Communication, Leadership

Relevant Coursework

- **Image synthesis:** Introduction to Diffusion Models, Computer Graphics, Realistic Image Synthesis, CV and ML for Computer Graphics
- **Image analysis:** Digital Image Processing, High-Level Computer Vision, Advanced Image Analysis
- **Image capture:** Image Acquisition Methods
- **Artificial Intelligence:** Machine Learning, Data Science, Neural Networks and Fuzzy Logic, Machine Learning in Cyber Security
- **Miscellaneous:** Data Structures and Algorithms, Operating Systems, Programming, Microprocessor Programming, Human-Computer Interaction