Mosam Dabhi

PH.D. CANDIDATE
Carnegie Mellon University, Pittsburgh, PA, USA
mosam@cmu.edu | mosamdabhi.github.io

Research Interests

AI systems with world-awarness.

Computer vision 3D Vision, Multi-view geometry, Physics [1–5]
AI / ML Optimization, SSL, Generative modeling, Graph [2–4, 6]
Graphics Implicit Functions, Differentiable Rendering [4]
Robotics Neuro-VLAs, SLAM, Control [7–11]

Education

Carnegie Mellon University *Pittsburgh, USA*
PH.D. IN ROBOTICS *Aug. '21 - Present*
- **Advisor**: Simon Lucey and László A Jeni

Carnegie Mellon University *Pittsburgh, USA*
M.S. IN ROBOTICS *Aug. '19 - May '21*
- **Advisor**: Simon Lucey
- Thesis: Multi-view NRSfM: Affordable Setup for High-Fidelity 3D Reconstruction

National Institute of Technology *Surat, India*
B.TECH. IN ELECTRONICS AND COMMUNICATION ENGINEERING *Aug. '13 - May '17*

Research Experiences

Carnegie Mellon University *Pittsburgh, USA*
PH.D. STUDENT *Aug. '19 - Present*
- Geometric Foundation Models to improve 3D vision, reasoning, and action by integrating geometry with foundational models.
- Developed models like 3D-LFM [3, 5, 7] advancing scalable 3D and lens understanding for real-world robot actions.
- Geometry-first, verifiable manipulation controller: 10x data-efficient vs VLA visuomotor policies with few-shot cross-robot transfer.
- Introduced RAT4D [4] that creates dense animatable 3D models from RGB videos, enabling novel applications like gen-AI video correction.
- MV-NRSfM [1], MBW [2] enables scalable auto-labeling with minimal supervision, while 3D-LFM lifts 2D to 3D landmarks with graph-based transformers.

Apple Inc. *Cupertino, US*
RESEARCH SCIENTIST INTERN (AI) *May '24 - Sep '24*
- LLMs, LVMs, and 3D (*Vision Pro team*).

Apple Inc. *Cupertino, US*
RESEARCH SCIENTIST INTERN (AI) *June '23 - Aug. '23*
- Multimodal time-series modeling, shaping the foundation for lightweight AI foundation models (*Vision Pro team*).

Apple Inc. *Cupertino, US*
RESEARCH SCIENTIST INTERN *May '22 - Aug. '22*
- Few-shot learning and Out-Of-Distribution adaptation algorithms (*Vision Pro team*).
- Auto-labeling in the wild, driving substantial financial savings by auto-generating 3D labels for computer vision applications .

Apple Inc. *Cupertino, US*
RESEARCH INTERN *May '21 - Aug. '21*
- Active learning and self-supervised learning strategies (*Vision Pro team*).

Apple Inc. *Cupertino, US*
RESEARCH INTERN *May '20 - Aug. '20*
- Foundational ML work on multi-view 3D geometry, enabling economical 3D groundtruth labels (*Vision Pro team*).

Carnegie Mellon University *Pittsburgh, US*

Research Assistant *May '17 - May. '19*

- **Advisor**: Nathan Michael
- Real-time robotic exploration policies for search and rescue, enabling advanced communication on extra-/sub-terrestrial surfaces [11].
- Robot policies for flights in GPS-denied terrains with accelerations over 12 m/s^2 [10].
- Experience-driven Model Predictive Control (E-MPC) for computationally constrained platforms [8, 12].
- Path planning policy in cluttered environments via mixed-integer programming [9].

Indian Institute of Science *Bangalore, India*

Research Intern *May '15 - Jul. '15*

- **Advisor**: Prasanta Kumar Ghosh
- Speech models for mobile devices via Markovian models.

Publications

[1] **Mosam Dabhi**, Chaoyang Wang, Kunal Saluja, László A Jeni, Ian Fasel, and Simon Lucey. High Fidelity 3D Reconstructions with Limited Physical Views. In *2021 International Conference on 3D Vision (3DV)*. IEEE, 2021.

[2] **Mosam Dabhi**, Chaoyang Wang, Tim Clifford, László A Jeni, Ian Fasel, and Simon Lucey. Multi-view Bootstrapping in the Wild. In *Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track*. NeurIPS, 2022.

[3] **Mosam Dabhi**, László A Jeni, and Simon Lucey. 3d-lfm: Lifting foundation model. *2024 Computer Vision and Pattern Recognition (CVPR)*, 2024.

[4] **Mosam Dabhi**, Simon Lucey, and László A Jeni. Rat4d: Rig and animate any object without templates in 4d. *2026 Winter Conference on Applications of Computer Vision (WACV)*, 2026.

[5] Mukai Yu, **Mosam Dabhi**, Liuyue Xie, Sebastian Scherer, and László A Jeni. Unified spherical frontend: Learning rotation-equivariant representations of spherical images from any camera, 2025.

[6] Liuyue Xie, George Z. Wei, Avik Kuthiala, Ce Zheng, Ananya Bal, **Mosam Dabhi**, Liting Wen, Taru Rustagi, Ethan Lai, Sushil Khyalia, Rohan Choudhury, Morteza Ziyadi, Xu Zhang, Hao Yang, and László A. Jeni. Maverix: Multimodal audio-visual evaluation reasoning index. In *Fortieth AAAI Conference on Artificial Intelligence (AAAI)*. AAAI, 2026. Accepted for Poster Presentation.

[7] **Mosam* Dabhi**, Peiqi* Yu, László A Jeni, and Changliu Liu. Zero-shot deformable manipulation across robots via geometric reasoning, Under submission, 2025.

[8] **Mosam Dabhi**, Alexander Spitzer, and Nathan Michael. Aggressive Flight Performance using Robust Experience-driven Predictive Control Strategies: Experimentation and Analysis. Technical Report CMU-RI-TR-19-08, Carnegie Mellon University, Pittsburgh, PA, June 2019.

[9] **Mosam Dabhi**, Vishnu Desaraju, and Nathan Michael. Planning Aggressive, Dynamically Feasible and Optimal Trajectories for Autonomous Vehicles in Cluttered Environments using Mixed Integer Programming. Technical report, Carnegie Mellon University, Pittsburgh, PA, 2016.

[10] Alex Spitzer, Xuning Yang, John Yao, Aditya Dhawale, Kshitij Goel, **Mosam Dabhi**, Matt Collins, Curtis Boirum, and Nathan Michael. Fast and agile vision-based flight with teleoperation and collision avoidance on a multirotor. In *International Symposium on Experimental Robotics*, pages 524–535. Springer, 2018.

[11] Wennie Tabib, Kshitij Goel, John Yao, **Mosam Dabhi**, Curtis Boirum, and Nathan Michael. Real-Time Information-Theoretic Exploration with Gaussian Mixture Model Maps. In *Robotics: Science and Systems*, 2019.

[12] **Mosam Dabhi**, Vishnu R Desaraju, and Nathan Michael. Evaluation of Explicit Experience-driven Predictive Control on a Computationally Constrained Platform. Technical report, Carnegie Mellon University, Pittsburgh, PA, June 2017.

Honors & Awards

’19 - ’23	**Apple Research Grant**, Apple Inc.
’17	**Research Scholarship**, FICCI, India
’16 - ’17	**Summer Scholar**, Robotics Institute Summer Scholar
’16	**Undergraduate thesis funding**, TEQIP Award, MHRD, Government of India
’25	**Swartz Center, CMU**, Innovation Commercialization Fellow Award

Academic Services

’22 - ’25	**Conference Paper Reviewer**, NeurIPS; CVPR; ICCV; ECCV; WACV
’21 - ’25	**Conference Paper Reviewer**, ICRA, IROS
’20	**Conference Paper Reviewer**, International Conference on Humanoid Robots
’23 - ’25	**Admissions Committee @ CMU**, Doctor of Philosophy (PhD), Master of Science, Robotics (MSR), Master of Science, Computer Vision (MSCV)
’22 - ’23	**M.S. in Robotics Thesis Committee** , Examinee: Heng Yu, Aarush Gupta
’17 - ’19	**Admissions & Admin. Committee**, Robotics Institute Summer Scholars, CMU (RISS)

Teaching Experiences

Carnegie Mellon University

Teaching Assistant

- **Spring 2022** : Robot Localization and Mapping with Prof. Michael Kaess
- **Fall 2022** : Geometry-Based Methods in Vision with Prof. Shubham Tulsiani

Relevant coursework

Carnegie Mellon University

Learning for 3D Vision (**A+**), Advanced Computer Vision (**A**), Geometry Methods in Vision (**A+**), Advanced Machine Learning (**A**), Convex Optimization (**A**), Robot Localization and Mapping / SLAM (**A+**), Math. Fundamentals for Robotics (**A**), Kinematics, Dynamics, and Controls (**A+**)

Proficient Skills

Programming languages

Primary: Python, C/C++, LaTeX, MATLAB

Secondary: CUDA, Lua, HTML, Javascript

Software libraries

Primary: PyTorch, TensorFlow, Blender, COLMAP

Secondary: OpenAIGym, Torch, Caffe, OpenCV, VLFeat, Pthread