

Xiaohan Fei

E-mail: hzhsfxh@gmail.com Website: <https://feixh.github.io>

EDUCATION	UNIVERSITY OF CALIFORNIA, LOS ANGELES Ph.D. in Computer Science Research Group: UCLA Vision Lab (http://vision.ucla.edu) GPA: 3.88/4.0 Thesis: Inertial-aided Visual Perception of Geometry and Semantics Advisor: Prof. Stefano Soatto ZHEJIANG UNIVERSITY, HANGZHOU, CHINA B.Eng. in Information and Communication Engineering Minor: Advanced honor Class of Engineering Education (ACEE), Chu-Kechen College GPA: 3.98/4.0(92.35/100) Thesis: Wide-baseline feature matching for panoramic images Thesis advisor: Prof. Zhiyu Xiang	Sept. 2014 - Sept. 2019 Sept. 2010 - June 2014
PROFESSIONAL INTERESTS	My professional interests lie at the intersection of computer vision, robotics, and machine learning. My current work at Amazon focuses on building foundation models for video generation including text-to-video and image-to-video. In the past, I worked on (i) Simultaneous Localization and Mapping (SLAM) and Visual-Inertial Odometry (VIO), (ii) depth prediction and completion, and (iii) semantic scene understanding.	
OPEN-SOURCE SOFTWARE	XIVO : a state-of-the-art localization and mapping software, forked by 100+ and starred by 700+ developers worldwide. For the rest, see my github page: https://github.com/feixh	
WORK EXPERIENCE	AMAZON AGI, BELLEVUE, WASHINGTON Principal Applied Scientist Overall tech lead for Amazon Nova Reel – Amazon’s video generation foundation model. AWS AI LABS, BELLEVUE, WASHINGTON Senior Applied Scientist Visual and multi-sensor localization and mapping, 3-D representation learning, and foundation model for video generation (Amazon Nova Reel) launched at AWS re:Invent 2024. AWS AI LABS, SEATTLE, WASHINGTON Applied Scientist Initiative for a cloud-based localization and mapping service. FACEBOOK REALITY LABS, REDMOND, WASHINGTON Research Scientist Surreal team, multi-sensor localization and mapping for AR/VR.	April 2025 - Present April 2022 - April 2025 April 2020 - March 2022 Sept. 2019 - March 2020
AWARDS & DISTINCTIONS	2019: Best Paper Award in Robot Vision, out of 2900 submissions, at ICRA 2019 2013: Meritorious Winner of Mathematical Contest in Modeling (top 15% of 6000 teams world-wide) 2012: National Scholarship (the highest honor for undergraduates in China)	
PUBLICATIONS * indicates equal contribution	[1] The Amazon Nova Family of Models: Technical Report and Model Card. In <i>arXiv</i> , 2024. [2] Chethan Parameshwara*, Alessandro Achille*, Matthew Trager, Xiaolong Li, Jiawei Mo, Ashwin Swaminathan, CJ Taylor, Dheera Venkatraman, Xiaohan Fei* , Stefano Soatto*. Towards visual foundational models of physical scenes. In <i>arXiv</i> , 2023. [3] Xiaohan Fei , Chethan Parameshwara, Jiawei Mo, Xiaolong Li, Ashwin Swaminathan, CJ Taylor, Paolo Favaro, Stefano Soatto. A Quantitative Evaluation of Score Distillation Sampling Based Text-to-3D. In <i>arXiv</i> , 2023. [4] Ziqi Lu, Jianbo Ye, Xiaohan Fei , Xiaolong Li, Jiawei Mo, Ashwin Swaminathan, Stefano Soatto. Fast sparse view guided nerf update for object reconfigurations. In <i>arXiv</i> , 2023. [5] Xiaolong Li, Jiawei Mo, Ying Wang, Chethan Parameshwara, Xiaohan Fei , Ashwin Swaminathan, CJ Taylor, Zhuowen Tu, Paolo Favaro, Stefano Soatto. Grounded compositional and diverse text-to-3d with pretrained multi-view diffusion model. In <i>arXiv</i> , 2023.	

- [6] **X. Fei**, H. Wang, X. Zeng, L. Cheong, J. Tighe. Single View Physical Distance Estimation using Human Pose. In *International Conference on Computer Vision (ICCV)*, 2021.
- [7] A. Wong, **X. Fei**, B. Hong, and S. Soatto. An Adaptive Framework For Learning Unsupervised Depth Completion. In *IEEE Robotics and Automation Letters (RA-L)*, 2021.
- [8] A. Wong*, **X. Fei***, and S. Soatto. Unsupervised Depth Completion from Visual-Inertial Odometry. In *International Conference on Robotics and Automation (ICRA)*, 2020. Also in *IEEE Robotics and Automation Letters (RA-L)*.
- [9] **X. Fei**, A. Wong, and S. Soatto. Geo-Supervised Visual Depth Prediction. In *International Conference on Robotics and Automation (ICRA)*, 2019. (**Best Paper Award in Robot Vision**) Also in *IEEE Robotics and Automation Letters (RA-L)*.
- [10] **X. Fei**, S. Soatto. Visual-Inertial Object Detection and Mapping. In *European Conference on Computer Vision (ECCV)*, 2018.
- [11] J. Dong*, **X. Fei***, and S. Soatto. Visual-Inertial-Semantic Scene Representation for Object Detection. In *Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [12] **X. Fei**, K. Tsotsos, and S. Soatto. A Simple Hierarchical Pooling Data Structure for Loop Closure. In *European Conference on Computer Vision (ECCV)*, 2016.

PROFESSIONAL SERVICES Reviewer of top computer vision (CVPR, ICCV, ECCV), robotics (ICRA, IROS), and artificial intelligence (AAAI) conferences.

RELEVANT COURSEWORK **University of California, Los Angeles:** Machine Perception (Prof. S. Soatto), Convex Optimization (Prof. L. Vandenberghe), Calculus of Variations (Prof. L. Vese), Vision as Bayesian Inference (Prof. A. Yuille), Applied Probability (Prof. Y. Wu), Theoretical Statistics (Prof. A. Amini), Numerical Analysis (Prof. J. Teran), Machine Learning Algorithm (Prof. M. Sarrafzadeh)
Zhejiang University: Computer Vision (Prof. Z. Xiang), Spectral Analysis of Signals (Prof. X. Gong), Information Theory (Prof. Z. Zhang), Mathematical Modeling (Prof. Q. Yang)

RELEVANT SKILLS **Programming Language:** Python, C++
Software Framework: Deep Learning (PyTorch, Megatron, TensorFlow), Vision and Robotics (OpenCV, ROS), Math & Optimization (Eigen, Ceres), Web (Flask, Three.js)