

# Xiaohan Fei

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

EDUCATION	<p>UNIVERSITY OF CALIFORNIA, LOS ANGELES Sept. 2014 - Sept. 2019 <b>Ph.D. in Computer Science</b> Research Group: UCLA Vision Lab (<a href="http://vision.ucla.edu">http://vision.ucla.edu</a>) GPA: 3.88/4.0 Thesis: Inertial-aided Visual Perception of Geometry and Semantics Advisor: Prof. Stefano Soatto</p> <p>ZHEJIANG UNIVERSITY, HANGZHOU, CHINA Sept. 2010 - June 2014 <b>B.Eng. in Information and Communication Engineering</b> 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</p>
RESEARCH EXPERIENCE	<p>AMAZON AGI, BELLEVUE, WASHINGTON April 2025 - Present <b>Principal Applied Scientist</b> I'm the science lead of Amazon Nova Reel – Amazon's foundation model for video generation launched at AWS re:Invent 2024.</p> <p>AWS AI LABS, BELLEVUE, WASHINGTON April 2022 - April 2025 <b>Senior Applied Scientist</b> I led a small team working on visual and multi-sensor localization and mapping and 3-D representation learning.</p> <p>AWS AI LABS, SEATTLE, WASHINGTON April 2020 - March 2022 <b>Applied Scientist</b> I conducted research in the field of computer vision and machine learning, and developed cloud-based AI services.</p> <p>FACEBOOK REALITY LABS, REDMOND, WASHINGTON Sept. 2019 - March 2020 <b>Research Scientist</b> I was a member of the Surreal team conducting research in computer vision for AR/VR.</p>
AWARDS & DISTINCTIONS	<p>2019: <b>Best Paper Award</b> in Robot Vision, out of 2900 submissions, at ICRA 2019 2013: <b>Meritorious Winner</b> of Mathematical Contest in Modeling (top 15% of 6000 teams worldwide) 2012: <b>National Scholarship</b> (the highest honor for undergraduates in China)</p>
PUBLICATIONS * indicates equal contribution	<p>[1] The Amazon Nova Family of Models: Technical Report and Model Card. In <i>arXiv</i>, 2024.</p> <p>[2] Chethan Parameshwara*, Alessandro Achille*, Matthew Trager, Xiaolong Li, Jiawei Mo, Ashwin Swaminathan, CJ Taylor, Dheera Venkatraman, <b>Xiaohan Fei*</b>, Stefano Soatto*. Towards visual foundational models of physical scenes. In <i>arXiv</i>, 2023.</p> <p>[3] <b>Xiaohan Fei</b>, 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.</p> <p>[4] Ziqi Lu, Jianbo Ye, <b>Xiaohan Fei</b>, Xiaolong Li, Jiawei Mo, Ashwin Swaminathan, Stefano Soatto. Fast sparse view guided nerf update for object reconfigurations. In <i>arXiv</i>, 2023.</p> <p>[5] Xiaolong Li, Jiawei Mo, Ying Wang, Chethan Parameshwara, <b>Xiaohan Fei</b>, 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.</p> <p>[6] <b>X. Fei</b>, H. Wang, X. Zeng, L. Cheong, J. Tighe. Single View Physical Distance Estimation using Human Pose. In <i>International Conference on Computer Vision (ICCV)</i>, 2021.</p> <p>[7] A. Wong, <b>X. Fei</b>, B. Hong, and S. Soatto. An Adaptive Framework For Learning Unsupervised Depth Completion. In <i>IEEE Robotics and Automation Letters (RA-L)</i>, 2021.</p> <p>[8] A. Wong*, <b>X. Fei*</b>, and S. Soatto. Unsupervised Depth Completion from Visual-Inertial Odometry. In <i>International Conference on Robotics and Automation (ICRA)</i>, 2020. Also in <i>IEEE Robotics and Automation Letters (RA-L)</i>.</p>

- [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:** Fluent (C++, Python), Make-do (JavaScript), Collecting dust (Android, OpenGL, MATLAB)  
**Software Framework:** Deep Learning (PyTorch, TensorFlow), Vision and Robotics (OpenCV, ROS), Math & Optimization (Eigen, Ceres), Web (Flask, Three.js)