

Mingfang Zhang

✉ mfzhang@iis.u-tokyo.ac.jp • 🏠 Homepage

Education

The University of Tokyo

Ph.D., supervised by Professor Yoichi Sato

M.Sc., supervised by Professor Yoichi Sato

Japan

2023.4–2026.3

2021.4–2023.3

Nanjing Univeristy

B.Sc. in Computer Science, Elite Class

China

2016.9–2020.8

PhD Research

I focused on computer vision and human activity understanding, specifically involving video and multi-view understanding, vision-language multimodal models, and human body perception.

Internship Experience

Woven by Toyota, Tokyo

Ongoing

2025

CyberAgent AI Lab, Tokyo

[B.2] *Egocentric Action-aware Inertial Localization in Point Clouds*

2024

Shanghai AI Lab, Shanghai

[B.1] *An Egocentric Vision-Language Model based Portable Real-time Smart Assistant*

[A.3] *EgoExoLearn: A Dataset for Bridging Asynchronous Ego- and Exo-View of Activities*

2023

Microsoft Research Asia, Beijing

[A.4] *Structural Multiplane Image: Bridging Neural View Synthesis and 3D Reconstruction*

2022

PCL Laboratory, Shenzhen

[C.4] *GazeOnce: Real-Time Multi-Person Gaze Estimation*

2021

Services and Awards

- Reviewer of CVPR, ICCV, ECCV, NeurIPS, ICML, ICLR, AAAI, ACMMM, BMVC, TCSVT, MVA
- JSPS Research Fellowship for Young Scientists DC2, 2025
- Program Committee of the Human-Autonomous Vehicle Interaction Workshop at WACV 2025
- UTokyo-IIS Research Collaboration Initiative Award 2024
- Honorable mention in essay competition at ICVSS 2024
- 1st place award of EgoTracks challenge in Ego4D at CVPR 2023
- “Stars of Tomorrow” award by Microsoft Research Asia, 2022
- Excellent Graduation Paper award by Nanjing University, 2020

Skills

Programming Languages/Tools: Python, PyTorch, Hugging Face, Git, Docker, Singularity, \LaTeX

Publication

A. Video and Multi-view Understanding

- [A.1] **Masked Video and Body-worn IMU Autoencoder for Egocentric Action Recognition**
Mingfang Zhang, Yifei Huang, Ruicong Liu, Yoichi Sato
(**ECCV**) The European Conference on Computer Vision, 2024 [pdf]
 - [A.2] **EgoExoLearn: A Dataset for Bridging Asynchronous Ego- and Exo-View of Activities**
(*co-first authors) *Yifei Huang*, Guo Chen*, Jilan Xu*, Mingfang Zhang*, ..., Yu Qiao*
(**CVPR**) IEEE Conference on Computer Vision and Pattern Recognition, 2024 [pdf]
 - [A.3] **Structural Multiplane Image: Bridging Neural View Synthesis and 3D Reconstruction**
Mingfang Zhang, Jinglu Wang, Xiao Li, Yifei Huang, Yoichi Sato, Yan Lu
(**CVPR**) IEEE Conference on Computer Vision and Pattern Recognition, 2023 [pdf]
 - [A.4] **Optical Flow in the Dark**
Mingfang Zhang, Yinqiang Zheng, Feng Lu
(**TPAMI**) IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021 [pdf]
 - [A.5] **Optical Flow in the Dark**
(*co-first authors) *Yinqiang Zheng*, Mingfang Zhang*, Feng Lu*
(**CVPR**) IEEE Conference on Computer Vision and Pattern Recognition, 2020 [pdf]
-

B. Vision-Language Multimodal Models

- [B.1] **Prompt-augmented Boundary Attentive Learning for Weakly Supervised Temporal Grounding**
Zhehao Zhu, Yifei Huang, Mingfang Zhang, Liangyang Ouyang, Yoichi Sato
(**TCSVT**) IEEE Transactions on Circuits and Systems for Video Technology, 2025
 - [B.2] **An Egocentric Vision-Language Model based Portable Real-time Smart Assistant**
Yifei Huang, Jilan Xu, Baoqi Pei, Yuping He, Guo Chen, Mingfang Zhang, ..., Limin Wang
(in submission) [pdf]
 - [B.3] **Egocentric Action-aware Inertial Localization in Point Clouds**
Mingfang Zhang, Ryo Yonetani, Yifei Huang, Liangyang Ouyang, Ruicong Liu, Yoichi Sato
(in submission) [pdf]
 - [B.4] **Humanity's Last Exam**
Long Phan, Alice Gatti, Ziwen Han, Nathaniel Li, ..., Mingfang Zhang, ..., Dan Hendrycks
ArXiv preprint [pdf]
-

C. Human Body Perception

- [C.1] **Event-Based Hand Pose Estimation with RGB Images Pre-Training**
Ruicong Liu, Takehiko Ohkawa, Tze Ho Elden Tse, Mingfang Zhang, Angela Yao, Yoichi Sato
(in submission)
- [C.2] **SiMHand: Mining Similar Hands for Large-Scale 3D Hand Pose Pre-training**
Nie Lin, Takehiko Ohkawa, Yifei Huang, Mingfang Zhang, ..., Yoichi Sato
(**ICLR**) International Conference on Learning Representations, 2025 [pdf]
- [C.3] **Single-to-Dual-View Adaptation for Egocentric 3D Hand Pose Estimation**
Ruicong Liu, Takehiko Ohkawa, Mingfang Zhang, Yoichi Sato
(**CVPR**) IEEE Conference on Computer Vision and Pattern Recognition, 2024 [pdf]
- [C.4] **GazeOnce: Real-Time Multi-Person Gaze Estimation**
Mingfang Zhang, Yunfei Liu, Feng Lu
(**CVPR**) IEEE Conference on Computer Vision and Pattern Recognition, 2022 [pdf]