Mingfang Zhang

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

Research Interests

My research interests lie in computer vision and egocentric human activity understanding, leveraging diverse data modalities including video, language, IMU, point cloud, human gaze, and hand pose.

Education

The University of Tokyo **Japan**
Ph.D., supervised by Professor Yoichi Sato *2023.4–2026.3*
M.Sc., supervised by Professor Yoichi Sato *2021.4–2023.3*

Nanjing Univeristy **China**
B.Sc. in Computer Science, Elite Class *2016.9–2020.8*

Internship Experience

CyberAgent AI Lab, Tokyo **2024**
Egocentric Inertial Navigation and Open-Vocabulary Action Recognition in the Point Clouds

- Proposed a novel paradigm for the inertial navigation task by exploiting the relationship between video, language, human motion, and point cloud data.
- Addressed the open-vocabulary action recognition task by integrating IMU sensor signals with point cloud data at dynamically estimated human positions.

Shanghai AI Lab, Shanghai **2023**
EgoBridge: A Dataset for Bridging Asynchronous First- and Third-Person View of Activities

- Introduced a large-scale dataset in which individuals record egocentric videos with gaze as they execute tasks guided by exo demonstration videos.
- Proposed a new benchmark, cross-view referenced skill assessment, aiming to rank the skill level of two ego videos with an exo video of expert demonstration as reference.

Microsoft Research Asia, Beijing **2022**
Structural Multiplane Image: Bridging Neural View Synthesis and 3D Reconstruction

- Presented a novel Structural MPI 3D representation, consisting of geometrically-faithful RGBα image layers to the 3D scene, for both neural view synthesis and 3D reconstruction.
- Proposed a network to construct the Structural MPI from posed images, where planar and non-planar regions are uniformly handled with approximations for geometries and light filed.

PCL Laboratory, Shenzhen **2021**
GazeOnce: Real-Time Multi-Person Gaze Estimation

- Proposed the first one-stage 3D gaze estimation method, estimating multi-user gaze simultaneously in a single image, and designed a projection-based self-supervised strategy for better accuracy.
- Released a new gaze dataset, enabling one-stage gaze estimation training and evaluation. It was generated by a sophisticated swap-gaze procedure with head pose matching.

Publication

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*
ArXiv preprint, 2025[pdf]

SiMHand: Mining Similar Hands for Large-Scale 3D Hand Pose Pre-training
Nie Lin, Takehiko Ohkawa, Yifei Huang, ***Mingfang Zhang****, ..., Yoichi Sato*
International Conference on Learning Representations (ICLR), 2025 [pdf]

Masked Video and Body-worn IMU Autoencoder for Egocentric Action Recognition
Mingfang Zhang*, Yifei Huang, Ruicong Liu, Yoichi Sato*
The European Conference on Computer Vision (ECCV), 2024 [pdf]

EgoExoLearn: A Dataset for Bridging Asynchronous Ego- and Exo-centric View of Activities
*(*co-first authors) Yifei Huang*, Guo Chen*, Jilan Xu*,* ***Mingfang Zhang*****, ..., Yu Qiao*
IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2024 [pdf]

Single-to-Dual-View Adaptation for Egocentric 3D Hand Pose Estimation
Ruicong Liu, Takehiko Ohkawa, ***Mingfang Zhang****, Yoichi Sato*
IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2024 [pdf]

Structural Multiplane Image: Bridging Neural View Synthesis and 3D Reconstruction
Mingfang Zhang*, Jinglu Wang, Xiao Li, Yifei Huang, Yoichi Sato, Yan Lu*
IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2023 [pdf]

GazeOnce: Real-Time Multi-Person Gaze Estimation
Mingfang Zhang*, Yunfei Liu, Feng Lu*
IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2022 [pdf]

Optical Flow in the Dark
Mingfang Zhang*, Yinqiang Zheng, Feng Lu*
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2021 [pdf]

Optical Flow in the Dark
*(*co-first authors) Yinqiang Zheng*,* ***Mingfang Zhang*****, Feng Lu (* co-first author)*
IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020 [pdf]

Services and Awards

- Reviewer of CVPR, ICCV, ECCV, NeurIPS, ICML, ICLR, AAAI, ACMMM, BMVC, TCSVT
- 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