

# Yu Rong

Email : rongyu9124@hotmail.com

Personal Homepage

Phone: (1)4122099653

Github & Google Scholar

## INTRODUCTION

---

I am currently a Research Engineer at (Meta) Reality Labs Research, Pittsburgh. Prior to that, I obtained the Ph.D. degree from The Chinese University of Hong Kong, Multimedia Laboratory, in September 2021, advised by Prof. Chen Change Loy and Prof. Xiao'ou Tang. My research interests include computer vision, computer graphics, and machine learning. Especially, I am interested in virtual humans, including 3D face, hand, and body reconstruction.

## EDUCATION

---

- **The Chinese University of Hong Kong** Hong Kong  
Doctor of Philosophy, Information Engineering Aug. 2017 – Sep. 2021
- **Tsinghua University** Beijing, China  
Bachelor of Engineering, Computer Science and Technology Aug. 2012 – Jun. 2016

## RESEARCH EXPERIENCE

---

- **Meta Reality Lab Research** Pittsburgh, PA, United States  
Research Engineer Mar. 2022 – Now
  - **Live Avatar Driving for AR.** Build the end-to-end pipeline for AR-based real-time avatar driving. The work includes raw data collecting, ground-truth generating from multi-view images, training 3D face tracker and expression encoder using collected data, and live system building and optimization. The live system takes inputs image from AR and generates headposes, 3D face geometries, and expression code to drive the avatar. The system can run at real-time performance and generalizes to different scenarios with various identities, headposes, expressions, lighting conditions and backgrounds.
- **Facebook AI Research** MPK, CA, United States  
Research Intern Jan. 2020 – May. 2020
  - **In-the-wild Whole Body 3D Motion Capture.** We use SMPL-X to represent whole body 3D motion including face, hands and body. Separate modules is adopt to predict independent body part first. Part motion predictions are then composed by the integration module to obtain unified whole body motion results. We design three different integration modules which trade off between accuracy and speed. Our system, both fast mode and slow-yet-precise mode, largely surpass the previous methods for in-the-wild scenarios. Demo, code and models are available at the project page.
- **The Chinese University of Hong Kong** Hong Kong  
Ph.D. Candidate Aug. 2017 – Jul. 2021
  - **Monocular 3D Interacting Hands Reconstruction.** We try to solve the challenging problem of reconstructing 3D interacting hands from monocular single RGB images. Simply extending the single-hand based methods would generate unnatural results with hands collided with each other. To resolve this, we design a factorized refinement strategy to progressively identify and refine the casual factors of the undesirable collisions. Our method can effectively remove the 49% generated collisions while improving the pose estimation accuracy by 9%.
  - **In-the-wild 3D Human Body Motion Capture.** SMPL is used to represent 3D humans. First, we thoroughly investigate the effectiveness of different annotations for in-the-wild images. We find that the model trained with dense correspondence such as DensePose achieves 92% performance of the model trained with paired 3D annotations. Second, we propose a new framework, PM-Net, in which 3D human prototype memory is leveraged to replace the single human prototype used by previous works. Using PM-Net, the models' prediction error on challenging in-the-wild samples are reduced by 12 mm.
  - **Pose-Robust Face Recognition** We hypothesize that discrepancy between frontal and profile faces could be bridged by an equivariant mapping. Guided by it, a novel deep residual equivariant mapping block is designed, which reduces the error rate of the baseline ResNet-50 model by 12.6% on the identification task of IJB-A dataset.

## PUBLICATIONS

---

1. Jingbo Wang, **Yu Rong**, *etc* “ Towards Diverse and Natural Scene-aware 3D Human Motion Synthesis ”, In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 2022.
2. **Yu Rong**, Jingbo Wang, Ziwei Liu, Chen Change Loy, “ Monocular 3D Reconstruction of Interacting Hands via Collision-Aware Factorized Refinements ”, In *International Conference on 3D Vision (3DV)* 2021.
3. Guanze Liu, **Yu Rong**, Lu Sheng, “ VoteHMR: Occlusion-Aware Voting Network for Robust 3D Human Mesh Recovery from Partial Point Clouds ” In *ACM Multimedia (MM)* 2021, **Oral**.
4. **Yu Rong**, Takaaki Shiratori, Hanbyul Joo, “FrankMocap: A Monocular 3D Whole-Body Pose Estimation System via Regression and Integration ”, In *International Conference on Computer Vision Workshops (ICCVW)* 2021.
5. **Yu Rong**, Ziwei Liu, Chen Change Loy, “Chasing the Tail in Monocular 3D Human Reconstruction with Prototype Memory”, In *IEEE Transaction on Image Processing (TIP)*.
6. **Yu Rong**, Ziwei Liu, Chen Change Loy *et al*, “Delving Deep into Hybrid Annotations for 3D Human Recovery in the Wild”, In *International Conference on Computer Vision (ICCV)* 2019.
7. Kaidi Cao\*, **Yu Rong**\*, Chen Change Loy *et al*, “Pose-Robust Face Recognition via Deep Residual Equivariant Mapping”, In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 2018. (\* indicates equal contribution)
8. Yong Tian, **Yu Rong**, Jiaxing Song *et al*, “Grouped text clustering using non-parametric Gaussian mixture experts”, In *Pacific Rim International Conference on Artificial Intelligence (PRICAI)* 2016.
9. **Yu Rong**, Xiaolin Hu, “ Fuzzy String Matching Using Sentence Embedding Algorithms ”, In *International Conference on Neural Information Processing (ICONIP)* 2016.

## ACADEMIC ACTIVITIES

---

- Co-organize a workshop on human sensing in computer vision at ICCV 2019.
- Serve as the reviewer for CVPR 2019~2022, ICCV 2019~2023, ECCV 2018~2022, ICLR 2022, and AAAI 2020/2021.
- Serve as the reviewer for TPAMI, IJCV, and TIP.

## SKILL

---

- Programming Languages: Python, C++, Java.
- Deep Learning Frameworks: Pytorch, LibTorch, Caffe.