

Center for Research in Computer Vision

UNIVERSITY OF CENTRAL FLORIDA

FINAL ORAL EXAMINATION

OF

Ce Zheng
B.S., University of Bridgeport, 2016
M.S., Tufts University, 2019

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

(COMPUTER SCIENCE)

14, NOV, 2023, 11:00 A.M.
HEC-101(A)

DISSERTATION COMMITTEE

Dr. Chen Chen, *Chairman*	*chen.chen@ucf.edu*
Dr. Mubarak Shah	*shah@crcv.ucf.edu*
Dr. Qian Lou	*qian.lou@ucf.edu*
Dr. Hao Zheng	*hao.zheng@ucf.edu*

DISSERTATION RESEARCH IMPACT

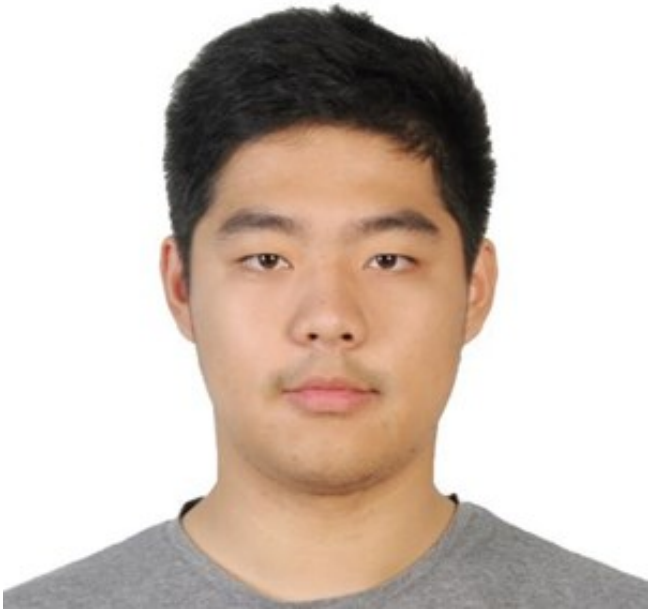
Understanding humans in visual content is fundamental for numerous computer vision applications. Extensive research has been conducted in the field of human pose estimation (HPE) to accurately locate joints and construct body representations from images and videos. Expanding on HPE, human mesh recovery (HMR) addresses the more complex task of estimating the 3D pose and shape of the entire human body. However, HPE and HMR come with notable challenges. Despite the progress made so far, the research community continues to strive for robust, accurate, and efficient solutions in HPE and HMR, advancing us closer to the ultimate goals in the field. This dissertation tackles various challenges in the domains of HPE and HMR. Our initial focus is on video-based HPE. Next, we address the heavy computational and memory burden associated with image-based and video-based HMR. These achievements provide practical and efficient solutions for real-world applications in HPE and HMR.

SELECTED PUBLICATIONS

1. **Deep learning-based human pose estimation: A survey,** Ce Zheng, Wenhan Wu, Chen Chen, Taojiannan Yang, Sijie Zhu, Ju Shen, Nasser Kehtarnavaz, and Mubarak Shah. *ACM Computing Surveys, 2023*.
2. **3D Human Pose Estimation with Spatial and Temporal Transformer,** Ce Zheng, Sijie Zhu, Matias Mendieta, Taojiannan Yang, Chen Chen, and Zhengming Ding, *in ICCV'2021*.
3. **A Lightweight Graph Transformer Network for Human Mesh Reconstruction from 2D Human Pose,** Ce Zheng , Matias Mendieta, Pu Wang, Aidong Lu, and Chen Chen, *in ACM Multimedia'2022*.
4. **FeatER: An Efficient Network for Human Reconstruction via Feature Map-Based TransformER,** Ce Zheng, Matias Mendieta, Taojiannan Yang, Guo-Jun Qi, and Chen Chen, *in CVPR'2023*.
5. **POTTER: Pooling Attention Transformer for Efficient Human Mesh Recovery.** Ce Zheng, Xianpeng Liu, Guo-Jun Qi, and Chen Chen, *in CVPR'2023*.
6. **PoseFormerV2: Exploring Frequency Domain for Efficient and Robust 3D Human Pose Estimation.** Qitao Zhao, Ce Zheng, Mengyuan Liu, Pichao Wang, and Chen Chen, *in CVPR'2023*.

DISSERTATION
RECONSTRUCTING 3D HUMANS FROM VISUAL DATA

It is crucial to understand humans in visual content for various computer vision applications. Human pose estimation (HPE) has been extensively studied to accurately locate joints and construct body representations from images and videos, facilitating advancements in human-computer interaction, action recognition, and motion analysis. Building upon HPE, human mesh recovery (HMR) takes on the more complex task of estimating the 3D pose and shape of the entire human body. HMR has emerged as a critical area of research with applications in digital human avatar modeling, AI coaching, virtual reality, and more. However, HPE and HMR present significant challenges including intricate body articulation, occlusion, depth ambiguity, and the limited availability of annotated 3D data. Despite notable progress, the research community continues to tackle these obstacles, striving to achieve robust, accurate, and efficient solutions in HPE and HMR, moving closer to the ultimate objectives in the field. This dissertation addresses several challenges in the field of HPE and HMR. We begin by focusing on video-based HPE, leveraging the transformer architecture to capture both spatial relationships between body joints and temporal correlations across frames. This approach allows us to effectively utilize the comprehensive connectivity and expressive power of transformers, enhancing the accuracy of pose estimation in video sequences. Building upon this, we then tackle the issue of heavy computational and memory burden in image-based HMR. Our proposed methods demonstrate superior performance while significantly reducing computational and memory requirements compared to existing state-of-the-art techniques. Furthermore, we extend our work to video-based HMR, a diffusion-based framework is proposed for reconstructing high-quality human mesh outputs given input video sequences. These advancements provide practical and efficient solutions that align with the demands of real-world applications in HPE and HMR. is to build systems that can perform robust, accurate, and efficient 3D human reconstruction.



Ce Zheng

2012-2016	B.S., University of Bridgeport, Bridgeport, CT.
2017-2019	M.S., Tufts University, Medford, MA
2019-2020	Research Assistant at Worcester Polytechnic Institute, Worcester, MA
2020-2021	Ph.D. student, University of North Carolina at Charlotte, Charlotte, NC
2022	Research Internship, InnoPeak Tech, WA
2020-2023	Ph.D., University of Central Florida, Orlando, FL