YANG ZHENG

2# Zijing Student Apartment ⋄ Tsinghua University ⋄ Beijing, 100084, P.R.China
(+86) 15659261781 ⋄ y-zheng18@mails.tsinghua.edu.cn ⋄ y.zheng18.tsinghua@gmail.com

EDUCATION

Tsinghua University, Beijing, P.R.China

Bachelor of Engineering in Automation *Aug 2018 - Jul, 2022 (expected)*

Bachelor of Psychology (second degree) *Aug 2019 - Jul, 2022 (expected)*

- **GPA: 3.92/4.0, Ranking: 1^{nd}/174**
- **2019 secretary** of the sports competition department in the student union of the Department of Automation

Core Courses

- **Mathematics**: Calculus A (4.0/4.0), Linear Algebra (4.0/4.0), Introduction to Complex Analysis (4.0/4.0), Probability and Statistics (4.0/4.0), Numerical Analysis and Algorithms (4.0/4.0).
- **CS**: Computer Languages and Programming (4.0/4.0), C++ Program Design and Training (4.0/4.0), Data Structure and Algorithms (4.0/4.0), Operating Systems (4.0/4.0), Fundamentals of Computer Graphics (4.0/4.0).

SCHOLARSHIPS & AWARDS

- **2020 Jiang Nanxiang Scholarship** (Highest scholarship for juniors in Tsinghua, $<$ **0.1%**)
- **2019 National Scholarship** (Highest scholarship given by the government of China, $<$ **0.1%**)
- **2019 Tsinghua Innovation Award of Science and Technology** (Awarded to undergraduate students with excellent research potentials, $<$**1%**)
- **2019 3^{rd} place** in the 21^{th} **Electronic Design Competition**, Tsinghua University
- **2018 2^{nd} place** in the 2^{nd} **Artificial Intelligence Challenge**, Tsinghua University

PUBLICATIONS & MANUSCRIPTS

1 **Yang Zheng**, Fei Xia, Tolga Birdal, Yanchao Yang, Yueqi Duan, Leonidas J. Guibas. 6D Camera Relocalization in Visually Ambiguous Extreme Environments. In submission to RA-L with ICRA2022.

2 **Yang Zheng***, Ruizhi Shao*, Yuxiang Zhang, Tao Yu, Zerong Zheng, Yebin Liu. DeepMultiCap: Performance Capture of Multiple Characters Using Sparse Multiview Cameras. Accepted by *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. [ArXiv][Project Page]

3 Yongming Rao, Wenliang Zhao, **Yang Zheng**, Yansong Tang, Jiwen Lu, Jie Zhou. VideoABC: A Real-World Video Dataset for Abductive Visual Commonsense Reasoning. [PDF]

RESEARCH INTEREST

Fields 3D Vision, Human Reconstruction, Video Understanding
Methods Deep Learning, Neural Networks

RESEARCH EXPERIENCES

Stanford University, CA, U.S.
Geometric Computing Group, Department of Computer Science *Apr, 2021 – Sep, 2021*
Research Assistant, Advisors: Profs. Leonidas J. Guibas
Project: 6D Camera Relocalization in Visually Ambiguous Extreme Environments

- Extended the scope of the camera relocalization to visually ambiguous extreme environments, and analyzed the challenges faced by current approaches through extensive experiments.
- Proposed a robust hierarchical localization framework by using temporal information to handle highly ambiguous situations and pose refinement to boost the accuracy while refining the 3D structure.
- Designed an environment-aware image enhancement module to optimally improve the image quality such that the downstream feature based reconstruction and localization losses are minimized.

Tsinghua University, Beijing, China
Broadband Network & Digital Media Lab, Department of Automation *Jul, 2020 – Mar, 2021*
Research Assistant, Advisors: Profs. Yebin Liu
Project: DeepMultiCap: Performance Capture of Multiple Characters Using Sparse Multiview Cameras

- Proposed a novel framework for high-fidelity multi-view reconstruction of multiple interacting characters, which achieved the state-of-the-art performance and could be extended to real world settings.
- Designed an efficient spatial-temporal attention-aware method to maximize reconstruction robustness and accuracy.
- Contributed and made public of a high-quality 3D human dataset, MultiHuman, containing 150 multi-person scans with detailed geometry and photorealistic texture.

Tsinghua University, Beijing, China
Intelligent Vision Group, Department of Automation *Jun, 2019 – Apr, 2020*
Research Assistant, Advisors: Profs. Jiwen Lu & Jie Zhou
Project: VideoABC: A Real-World Video Dataset for Abductive Visual Commonsense Reasoning

- Conceptualized a new task of abductive visual commonsense reasoning that requires the vision systems to infer the most plausible sequence of steps given two observations that describe the initial configuration and desired goal.
- Introduced a real-world video dataset for abductive visual commonsense reasoning task, which consists of 46,354 clips derived from 11,827 videos.
- Proposed a Hierarchical Dual Reasoning Network (HDRNet) to capture the long-term dependencies among steps and observations, and thus significantly outperformed baseline models.

PROGRAMMING SKILLS

Proficient	Python, Pytorch, C/C++, Latex, Markdown
Familiar	TensorFlow, MATLAB, Qt, Linux, etc.

LANGUAGE SKILLS

TOEFL iBT	104/120 (Reading 30, Listening 24, Speaking 24, Writing 26)
GRE	330/340+3.5/6.0 (Verbal 160, Quantitative 170, Analytical Writing 3.5)