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: 3.91/4.0, Ranking: 2^{nd}/173**
- **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***, 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]

2 Yongming Rao, Guangyi Chen, **Yang Zheng**, Benlin Liu, Jiwen Lu, Jie Zhou. Variational Discriminative Aggeration for Video Person Re-identification. [PDF]

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

Tsinghua University, Beijing, China
Broadband Network & Digital Media Lab, Department of Automation *Apr, 2020 – Dec, 2020*
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 scenes, which achieved the state-of-the-art performance and could be extended to real world settings.
- Leveraged the human shape and pose prior SMPL model estimated from multi-view images to improve the robustness of reconstruction under heavy occlusions.
- Designed an efficient attention-aware module to perform multi-view feature fusion with an informative view representation based on SMPL, which effectively guided the network to reserve fine-grained details.

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 including TSN, R3D, etc.

Project: Variational Discriminative Aggeration for Video Person Re-identification

- Proposed a variational discriminative aggregation method to explicitly model the structure of the feature space of the image sequence.
- Learned a motion-aware representation which extracted coarse-grained appearance and movement information by a two-stream network.
- Achieved the state-of-the-art performance on several benchmarks including Mars, DukeMTMC-VideoReID, and iLIDS-VID.

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)