CHENGZHI CAO

Email: chengzhicao@mail.ustc.edu.cn | Homepage
Hefei, China

EDUCATION

University of Science and Technology of China — **Project 985 & 211, Anhui, China**
School of Information Science and Technology — 2021.9 - Present
Master of Electronic Information

South China University of Technology — **Project 985 & 211, Guangdong, China**
Bachelor of Electrical Engineering and Automation — 2017.9 - 2021.6
Overall GPA: **3.86/4.00**

HONORS

- **National Scholarship (top1%, highest scholarship from Ministry of Education of China)** **2018**
- **National Scholarship (top1%, highest scholarship from Ministry of Education of China)** **2019**
- **National Scholarship (top1%, highest scholarship from Ministry of Education of China)** **2020**
- **Top Undergraduate Student Award (highest honor of undergraduates, only 10 winners)** **2021**
- **National Scholarship (top1%, highest scholarship from Ministry of Education of China)** **2023**

RESEARCH INTERESTS

Video restoration (deblurring and super-resolution) and bio-inspired intelligence

PUBLICATIONS

[1] Discovering Intrinsic Spatial-Temporal Logic Rules to Explain Human Actions
Chengzhi Cao, Chao Yang, Ruimao Zhang, Shuang Li*
Conference on Neural Information Processing Systems (NeurIPS Accept), Sep. 2023. [Paper]

[2] Event-guided Person Re-Identification via Sparse-Dense Complementary Learning
Chenzhi Cao, Xueyang Fu*, Hongjian Liu, Yukun Huang, Kunyu Wang, Jiebo Luo, Zheng-jun Zha
IEEE Conference on Computer Vision and Pattern Recognition (CVPR Accept). Mar. 2023 [Paper] [Code]

[3] Event-guided Video Restoration with Spiking-Convolutional Architecture
Chenzhi Cao, Xueyang Fu*, Yurui Zhu, Zhijing Sun, Zheng-jun Zha
IEEE Transactions on Neural Networks and Learning Systems (TNNLS Accept), Nov. 2023 [Paper]

[4] Event-driven Video Deblurring via Spatio-Temporal Relation-Aware Network
Chenzhi Cao, Xueyang Fu*, Yurui Zhu, Gege Shi, Zheng-jun Zha
International Joint Conference on Artificial Intelligence (IJCAI Long Oral Accept). Apr. 2022 [Paper] [Code]

[5] Generalized UAV Object Detection via Frequency Domain Disentanglement
Kunyu Wang, Xueyang Fu*, Hongjian Liu, Yukun Huang, **Chengzhi Cao**, Gege Shi, Zheng-jun Zha
IEEE Conference on Computer Vision and Pattern Recognition (CVPR Accept). Mar. 2023 [Paper]

[6] Single Image Shadow Detection via Complementary Mechanism
Yurui Zhu, Xueyang Fu*, **Chengzhi Cao**, Xi Wang, Qibin Sun, Zheng-jun Zha
Proceedings of the ACM International Conference on Multimedia (ACM MM Accept). Jun. 2022 [Paper][Code]

RESEARCH EXPERIENCE

Bio-inspired Video Processing with Guidance of Events — University of Science and Technology of China
Advisors: **Prof.Xueyang Fu** and **Zhengjun Zha**

Video Deblurring

- Propose a spatial-temporal relation-aware network for event-driven video deblurring, and achieve better performance through fusing features of frames and events properly.
- Propose a spiking neural temporal memory module by capturing long-term relations of event sequences.
- Extract the spatial correlation between frames and events to exploit the complementary information from them.
- Extensive experiments show that our method achieves the SOTA performance on banchmark of video deblurring dataset, including GoPro and HQF.

Video Person Identification

- The first event-guided solution to tackle the video-based Re-ID task.

- Propose a sparse-dense complementary learning network to fully utilize the sparse events and dense frames simultaneously to enhance identity representation learning in degraded conditions.
- Design a deformable spiking neural network to suit the sparse characteristics of event streams, which greatly utilizes the spatial consistency of events to provide motion information for dense RGB frames.

Spatial-Temporal Logic Learning to Explain Human Actions Chinese University of Hong Kong(Shenzhen)
Advisors: **Prof.Shuang Li** and **Ruimao Zhang**

Discovering Intrinsic Spatial-Temporal Logic Rules

- Propose a tractable and differentiable algorithm that can jointly learn the rule content and model parameters from observational data.
- The overall procedure is an expectation-maximization algorithm, where we treat the rule set as latent variables.
- In the E-step, the posterior distribution over the latent rule set is evaluated. In the M-step, the model parameters are optimized by maximizing the expected log-likelihood with respect to the current posterior.
- We demonstrated the promising performance of our model in terms of human action prediction and explanation on two interesting real datasets.

Enhancing Human-AI Collaboration Through Logic-guided Reasoning

- Utilize logical reasoning by generating some generalized knowledge about the agent's goal in the new environment to achieve robust social perception and human-AI collaboration.
- Employ a hierarchical reinforcement learning model with Theory of Mind (ToM) to plan robot actions for assisting humans.
- Analyze the movement sequence of agents in common household activities and obtain novel insights from generated logic rules.

Low-quality Video Detection and Filtering Tencent, Shenzhen
Advisors: **Prof.Kwok Wai Hung and Simon Lui**

- Design a multi-scale structure to take advantage of inter-scale correlation for video super-resolution in QQ Music, optimize it by removing unnecessary modules to simplify the network architecture.
- Smoke Detection. Identify the key points of multiple faces, and use them to locate the user's mouth; calculate the matching parameters between the cigarette and mouth. This project is now under implementation.
- **Patent Application Number:** 2021111598150, 2021109153727, 2021107976749

COMPETITIONS

Third Prize: CVPR Workshop MIPI-challenge - RGBW Joint Fusion and Denoise (CVPR Workshop)

Third Prize: CVPR Workshop NTIRE 2023 - HR Depth from Images of Transparent Surfaces (CVPR Workshop)

TEACHING EXPERIENCE

University of Science and Technology of China
Artificial Intelligence Theory (210707.01)
Autumn 2022 with Zheng-jun Zha, Zhiwei Xiong and Zilei Wang

SERVICES

Program Committee/External Reviewer: ICCV 2023, ACM MM 2023

SKILLS

Interests: behavior modeling, logic learning and low-level vision (deblurring, super-resolution and deraining)

Programming: Python (Pytorch, Opencv), LATEX, Matlab, Microsoft Office

Mathematics: Linear Algebra, Vector Calculus, Probability and Statistic

English: GRE 328 (V158+Q170+AW3.5) TOEFL 102 (R30+L25+S21+W26)