**Chengzhi Cao**

Email: [chengzhicao@mail.ustc.edu.cn](mailto:chengzhicao@mail.ustc.edu.cn) | [Homepage](https://chengzhi-cao.github.io/)

Hefei, China

Education

**South China University of Technology Project 985 & 211, Guangdong, China**

Bachelor of Technology in Electrical Engineering and Automation 2017.9 - 2021.6

Overall GPA: 3.85/4.0

**University of Science and Technology of China Project 985 & 211, Anhui, China**

School of Information Science and Technology 2021.9 - Present

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**

• First Prize in RoboMaster University League 2020

• First Prize in The Chinese Mathematics Competitions 2021

Research INTERESTs

Video Restoration (deblurring and super-resolution), bayes learning and bio-inspired Intelligence

Publications

### **[1] [Event-guided Person Re-Identification via Sparse-Dense Complementary Learning](https://www.ijcai.org/proceedings/2022/0112.pdf)**

**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](https://openaccess.thecvf.com/content/CVPR2023/papers/Cao_Event-Guided_Person_Re-Identification_via_Sparse-Dense_Complementary_Learning_CVPR_2023_paper.pdf)] [[Code](https://github.com/Chengzhi-Cao/SDCL)]

### **[2] [Event-driven Video Deblurring via Spatio-Temporal Relation-Aware Network](https://www.ijcai.org/proceedings/2022/0112.pdf)**

**Chenzhi Cao**, Xueyang Fu\*, Yurui Zhu, Gege Shi, Zheng-jun Zha

**IJCAI (Long Oral Accept, Acceptance Rate<3.75%)**. Apr. 2022 [[Paper](https://www.ijcai.org/proceedings/2022/0112.pdf)] [[Code](https://github.com/Chengzhi-Cao/STRA)]

### **[3] [Generalized UAV Object Detection via Frequency Domain Disentanglement](https://www.ijcai.org/proceedings/2022/0112.pdf)**

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

**[4] 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](https://dl.acm.org/doi/pdf/10.1145/3503161.3547904)][[Code](https://github.com/zhuyr97/SDCM)]

**[5] Discovering Intrinsic Spatial-Temporal Logic Rules to Explain Human Actions**

**Chengzhi Cao**, Chao Yang, Shuang Li\*

**Conference on Neural Information Processing Systems (NeurIPS Under Review)**, May. 2023.

**[6] 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 Under Review)**, Sep. 2022

**[7] Generalized Cryo-Electron Tomography Image Denoising via Frequency Domain Mining**

**Chengzhi Cao**, Tianyang Wang, Zhaoxin Fan, Min Xu\*

**International Conference on Computer Vision (ICCV Under Review)**, May. 2023.

Research Experience

**Bio-inspired Video Processing with Guidance of Events** University of Science and Technology of China

Sep. 2021 - Mar. 2022

Advisors: **Prof.[Xueyang Fu](https://xueyangfu.github.io/) and [Zhengjun Zha](https://research.com/u/zheng-jun-zha)**

**Video Deblurring**

• Propose a spatio-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 eventguided 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.

**Generalized Cryo-Electron Tomography Image Denoising** Carnegie Mellon University

Oct. 2022 - Mar. 2023

Advisors: **Prof.[Min Xu](https://xulabs.github.io/min-xu/)**

• Design a novel framework for unsupervised Cryo-Electron Tomography Image denoising, leveraging domain adaptation and Fourier transform.

• Develop a Fourier style mining generator to produce high-quality real-like images by converting images into frequency domain and then identifying domain-variant frequency components.

• Propose a cross distillation and refinement mechanism to learn domain-invariant representations by distribution alignment, narrowing the distribution gap between synthetic and real data domains.

**Spatial-Temporal Logic Learning to Explain Human Actions** Chinese University of Hong Kong

May. 2022 - Present

Advisors: **Prof.[Shuang Li](https://shuangli01.github.io/)**

• 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 64 our model in terms of human action prediction and explanation on two interesting real datasets.

**Low-quality Video Detection and Filtering** Tencent, Shenzhen

March. 2021 - July. 2021

Advisors: **Prof.[Kwok Wai Hung](https://scholar.google.nl/citations?user=61DydQUAAAAJ&hl=zh-CN&oi=ao) and [Simon Lui](https://scholar.google.nl/citations?hl=zh-CN&user=zP8lA0UAAAAJ)**

• 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](https://mipi-challenge.org/MIPI2023/" \l "awards) - RGBW Joint Fusion and Denoise

**Third Prize:** [CVPR Workshop NTIRE 2023](https://codalab.lisn.upsaclay.fr/competitions/10494" \l "results) - HR Depth from Images of Specular and Transparent Surfaces

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 (V158+Q170+AW3.5)