

Sizhuo Ma

sizhuoma@gmail.com

Personal website • Google Scholar • LinkedIn

SUMMARY

- Research interests: image/video restoration and enhancement, generative AI, computer vision, machine learning.
- Strong knowledge in low-level vision, multimodal LLMs, diffusion models, and computational imaging.
- Extensive experience in the design, training, performance optimization and on-device deployment of deep neural networks.
- Publications in top-tier conferences (SIGGRAPH, ECCV, CVPR, MobiCom) and journals (TOG, TPAMI, IJCV).

EDUCATION

University of Wisconsin–Madison Dec. 2016 – Jan. 2022

Ph.D. in Computer Sciences

Thesis: Resolving Motion with Single-Photon Cameras

University of Wisconsin–Madison Aug. 2014 – Dec. 2016

M.S. in Computer Sciences

GPA: 3.92/4.00

Shanghai Jiao Tong University, China Sep. 2010 – Jul. 2014

B.S. in Computer Science and Engineering

GPA: 90.3/100

WORK EXPERIENCE

Snap Inc. Feb. 2022 – Present

Senior Research Scientist

- Develop and deploy spatial/temporal frame upsamplers for **efficient video generation**.
- Design visual token pruning methods that reduce the inference cost of **vision-language models**.
- Build server-side and on-device **video super-resolution** models that improve video quality and user engagement.
- Train video **engagement prediction** models for large-scale content understanding and recommendation.
- Develop on-device **face image quality assessment** models for real-time visual quality evaluation.
- Create **computational imaging and display** technologies that improve user experience.
- **Mentor** research interns and publish research in top-tier computer vision and graphics venues.

WISION Lab, University of Wisconsin–Madison May. 2016 – Jan. 2022

Graduate Research Assistant (Advisor: Prof. Mohit Gupta)

- Developed novel solutions to motion-related computer vision problems (*e.g.*, scene flow, burst photography) with computational camera designs (*e.g.*, light field, structured light, single-photon cameras).

Snap Inc. May. 2020 – Aug. 2020

Research Intern, Computational Imaging Team (Supervisor: Shree Nayar)

- Developed computational methods that significantly improve the decoding distance of QR codes (4×).

Living Environments Lab, University of Wisconsin–Madison Jan. 2016 – May. 2016

Graduate Research Assistant (Advisor: Prof. Kevin Ponto)

- Built mobile AR prototypes with Unity, OpenCV and Caffe.

SELECTED PUBLICATIONS

Full publication list: [Google Scholar](#).

- [1] Guohao Sun, Yufei Wang, **Sizhuo Ma**, Yuege Xie, Yuting Cheng, Zhiqiang Tao, Jian Wang. IF-Prune: Information-Flow Guided Token Pruning for Efficient Vision-Language Models. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2026)*
- [2] Zhihang Zhong, Gurunandan Krishnan, Wei Wang, Xiao Sun, Yu Qiao, **Sizhuo Ma***, Jian Wang*. Velocity Disambiguation for Video Frame Interpolation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI 2026)*
- [3] Aryan Garg, **Sizhuo Ma**, Mohit Gupta. gQIR: Generative Quanta Image Reconstruction. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2026)*
- [4] Jinpei Guo, Yifei Ji, Zheng Chen, Yufei Wang, **Sizhuo Ma**, Yong Guo, Yulun Zhang, Jian Wang. Towards Redundancy Reduction in Diffusion Models for Efficient Video Super-Resolution. *arXiv, 2025*
- [5] **Sizhuo Ma**, Wei-Ting Chen, Qiang Gao, Jian Wang, Chris Wei Zhou, Wei Sun, Weixia Zhang, Linhan Cao, *et al.* VQualA 2025 Challenge on Face Image Quality Assessment: Methods and Results. *ICCV Workshop 2025*
- [6] Howard Zhang+, Yuval Alaluf+, **Sizhuo Ma**, Achuta Kadambi, Jian Wang*, Kfir Aberman*. InstantRestore: Single-Step

Personalized Face Restoration with Shared-Image Attention. *SIGGRAPH 2025*

- [7] Pradyumna Chari, **Sizhuo Ma**, Daniil Ostashev, Achuta Kadambi, Gurunandan Krishnan, Jian Wang*, Kfir Aberman. Personalized Restoration via Dual-Pivot Tuning. *IEEE Transactions on Image Processing (TIP 2025)*
- [8] Baiang Li, **Sizhuo Ma**, Yanhong Zeng, Xiaogang Xu, Youqing Fang, Zhao Zhang, Jian Wang*+, Kai Chen*. Exposure-Limited Image Enhancement with Generative Diffusion Prior. *International Conference on Computational Photography (ICCP 2025)*
- [9] Aditya Arora, Zhengzhong Tu+, Yufei Wang+, Ruizheng Bai, Jian Wang*, **Sizhuo Ma***. GuideSR: Rethinking Guidance for One-Step High-Fidelity Diffusion-Based Super-Resolution. *arXiv, 2025*
- [10] Wei-Ting Chen+, Vong Yu Jiet+, Yi-Tsung Lee, Qiang Gao, Sy-Yen Kuo, **Sizhuo Ma***, Jian Wang*. DiffVQA: Video Quality Assessment Using Diffusion Feature Extractor. *arXiv, 2025*
- [11] Jian Wang, **Sizhuo Ma**, Karl Bayer, Yi Zhang, Peihao Wang, Bing Zhou, Shree Nayar, Gurunandan Krishnan. Perspective-Aligned AR Mirror with Under-Display Camera. *SIGGRAPH Asia 2024 (Journal, Best Paper Award)*
- [12] Zhihang Zhong, Gurunandan Krishnan, Xiao Sun, Yu Qiao, **Sizhuo Ma***, Jian Wang*. Clearer Frames, Anytime: Resolving Velocity Ambiguity in Video Frame Interpolation. *European Conference on Computer Vision (ECCV 2024, Oral)*
- [13] Dasong Li, Wenjie Li, Baili Lu, Hongsheng Li, **Sizhuo Ma**, Gurunandan Krishnan, Jian Wang*. Delving Deep into Engagement Prediction of Short Videos. *European Conference on Computer Vision (ECCV 2024)*
- [14] Wei-Ting Chen, Gurunandan Krishnan, Qiang Gao, Sy-Yen Kuo, **Sizhuo Ma***, Jian Wang*+. DSL-FIQA: Assessing Facial Image Quality via Dual-Set Degradation Learning and Landmark-Guided Transformer. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2024)*
- [15] Wei-Ting Chen, Yu-Jiet Vong, Sy-Yen Kuo, **Sizhuo Ma***, Jian Wang*. RobustSAM: Segment Anything Robustly on Degraded Images. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2024, Highlight)*
- [16] **Sizhuo Ma**, Jian Wang, Wenzheng Chen, Suman Banerjee, Mohit Gupta, Shree Nayar. QfaR: Location-Guided Scanning of Visual Codes from Long Distances. *International Conference on Mobile Computing and Networking (MobiCom 2023)*
- [17] **Sizhuo Ma**, Shantanu Gupta, Arin C. Ulku, Claudio Bruschini, Edoardo Charbon, Mohit Gupta. Quanta Burst Photography. *SIGGRAPH 2020 / ACM Transactions on Graphics*

PATENTS

Real-time Selfie Perspective Undistortion on Mobiles by Im2Im Translation

Inventors: Jian Wang, Haiwei Chen, Sizhuo Ma, Gurunandan Krishnan Gorumkonda, US 20260004410 A1

Object Scale Utilizing Away-Facing Images

Inventors: Gurunandan Krishnan Gorumkonda, Rui Yu, Yicheng Wu, Jian Wang, Sizhuo Ma, US 20250238994 A1

AR Mirror

Inventors: Shree K. Nayar, Gurunandan Krishnan Gorumkonda, Jian Wang, Bing Zhou, Sizhuo Ma, Karl Bayer, Yicheng Wu, US 20240355239 A1

Energy-Efficient Adaptive 3D Sensing

Inventors: Jian Wang, Sizhuo Ma, Brevin Tilmon, Yicheng Wu, Gurunandan Krishnan Gorumkonda, Ramzi Zahreddine, Georgios Evangelidis, **Patent granted**, US Patent 12001024

Location-Guided Scanning of Visual Codes

Inventors: Sizhuo Ma, Jian Wang, Mohit Gupta, Shree K. Nayar, US 2022/0262089 A1

Systems, Methods, and Media for High Dynamic Range Quanta Burst Imaging

Inventors: Mohit Gupta, Sizhuo Ma, **Patent granted**, US Patent 11170549 and 11721054

Systems, Methods, and Media for Determining Object Motion in Three Dimensions from Light Field Image Data

Inventors: Mohit Gupta, Sizhuo Ma, Brandon Smith, **Patent granted**, US Patent 10706564

HONORS AND AWARDS

2022	Outstanding Graduate-Student Research Award, UW–Madison Computer Sciences Department
2020	Snap Research Fellowship
2012	Shanghai Municipal Scholarship
2011 – 2012	SJTU Academic Excellence Scholarship

SKILLS

Programming Languages: Python, MATLAB, C, C++

Tools/Libraries: OpenCV, PyTorch, CUDA, TensorRT, CoreML

Languages: English (Proficient), Chinese (Native), Japanese (JLPT N1)