Yuhao Zhu

CONTACT INFORMATION

Associate Professor
University of Rochester
3501 Wegmans Hall
Rochester, NY 14627

yzhu@rochester.edu
https://yuhaozhu.com/
https://horizon-lab.org/

RESEARCH INTERESTS

I do research at the intersection of imaging (optics and image sensors), human visual perception (psychophysical and computational modeling), and computer systems (computer architecture and programming systems) with the goal of fostering a sustainable and ecological way of human-computer integration. My current research has three complementary themes:

- modeling the neural and psychophysical basis of human perception and cognition for enhancing computing and imaging systems
- developing imaging and computer systems to augment human perception and cognition
- building modern computational tools (e.g., generative AI) to explore new forms of visual arts and to analyze art history

EDUCATION

The University of Texas at Austin
Ph.D., Electrical and Computer Engineering — *May 2017*
M.S.E., Electrical and Computer Engineering — *May 2015*
Dissertation: Energy-Efficient Mobile Web Computing

Beihang University, Beijing, China
B.S., Computer Science and Engineering — *June 2010*

ACADEMIC HONORS

- **2023, Bridging Fellowship, University of Rochester**
- **2023, University Research Award, University of Rochester**
- **2023, ISCA Hall of Frame**
- **2022, MICRO Hall of Frame**
- **2020, NSF CAREER Award**
- **2019, University Research Award, University of Rochester**
- **2018, ACM SIGARCH – IEEE-CS TCCA Outstanding Dissertation Award, Honorable Mention**
- **2017, Google Faculty Research Award**
- **2016, Google Ph.D. Fellowship**
- **2011, Microelectronics and Computer Development Fellowship, UT Austin**

PUBLICATION AWARDS

- **2022, IEEE Micro Top Picks of Computer Architecture, Honorable Mention**
  "*ANT: Exploiting Adaptive Numerical Data Type for Low-bit DNN Quantization*"
  Originally published at MICRO 2022

- **2022, IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR), Best Paper Honorable Mention**
  "*Real-Time Gaze Tracking with Event-Driven Eye Segmentation*"
- **2022, IEEE International Symposium on High-Performance Computer Architecture (HPCA), Best Paper Nominee**
  "*S2TA: Exploiting Structured Sparsity for Energy-Efficient Mobile CNN Acceleration*"
- **2022, Kostas Pantazos Memorial Award for Outstanding Paper in Visualization and Data Analysis, Society for Imaging Science and Technology**
  "*Digital Reconstruction of Elmina Castle for Mobile Virtual Reality via Point-based Detail Transfer*"
- **2020, International Conference on Parallel Architectures and Compilation Techniques (PACT), Best Paper Nominee**
  "*Low-Latency Proactive Continuous Vision*"
- **2019, IEEE Micro Top Picks of Computer Architecture**
  "*Energy-Efficient Video Processing for Virtual Reality*"
  Originally published at ISCA 2019
- **2019, IEEE/ACM International Symposium on Microarchitecture (MICRO), Best Paper Nominee**
  "*Tigris: Architecture and Algorithms for 3D Perception in Point Clouds*"
- **2018, IEEE Micro Top Picks of Computer Architecture, Honorable Mention**
  "*Euphrates: Algorithm-SoC Co-Design for Low-Power Mobile Continuous Vision*"
  Originally published at ISCA 2018
- **2015, IEEE Micro Top Picks of Computer Architecture, Honorable Mention**
  "*Microarchitectural Implications of Event-driven Server-side Web Applications*"
  Originally published at MICRO 2015
- **2014, Best of Computer Architecture Letters Award**
  "*Exploiting Webpage Characteristics for Energy-Efficient Mobile Web Browsing*"
- **2010, Outstanding Undergraduate Thesis Award, Beihang University**
  "*Distributed Time, Conservative Parallel Logic Simulation on GPUs*"

MENTORING AWARDS

- **2020, CRA Outstanding Undergraduate Researcher, Honorable Mention**
  2019, ACM Student Research Competition (ASPLOS 2019), Gold Medalist
  Sam Triest (Advisee)
- **2019, CRA Outstanding Undergraduate Researcher, Honorable Mention**
  2019, A. Richard Newton Young Student Fellow (DAC 2019)
  Qiuyue Sun (Advisee)
- **2020, CRA Outstanding Undergraduate Researcher, Honorable Mention**
  Sifan Ye (Advisee)

PROFESSIONAL POSITIONS

University of Rochester

Associate Professor

- Department of Computer Science — July 2024 - now
- Department of Brain and Cognitive Sciences — July 2024 - now

Assistant Professor

- Department of Computer Science — Jan. 2018 - July 2024
- Department of Brain and Cognitive Sciences — Sept. 2023 - July 2024

Affiliated Faculty
- Goergen Institute for Data Science — June 2018 - now
- Center for Visual Science — Sept. 2023 - now

Arm Research
Visiting Researcher — July 2017 - Dec. 2017

Harvard University
Research Fellow — Aug. 2016 - May 2017

The University of Texas at Austin
Graduate Researcher — Aug. 2010 - May 2017
Teaching Assistant — Fall 2010, Spring 2011, Spring 2014

Google Inc.
SWE Intern (Lead of Flow API of the Catapult project) — Summer 2015

AMD Research Lab
Research Intern — Summer 2012, Summer 2013

STMicroelectronics
Co-op Engineer — Summer 2011

Tsinghua University
Undergraduate Researcher — June 2009 - May 2010

PUBLICATIONS

Journal Articles

- Shuang Wu, Bo Yu, Shaoshan Liu, Yuhao Zhu
  Autonomy 2.0: The Quest for Economies of Scale
  Communications of the ACM (to appear)

- Zishen Wan, Yiming Gan, Bo Yu, Shaoshan Liu, Arijit Raychowdhury, Yuhao Zhu
  The Vulnerability-Adaptive Protection Paradigm Toward Reliable Autonomous Machines
  Communications of the ACM, 67.9 (2024): 66-77

- Yu Feng, Weikai Lin, Zihan Liu, Jingwen Leng, Minyi Guo, Han Zhao, Xiaofeng Hou, Jieru Zhao, Yuhao Zhu
  Potamoi: Accelerating Neural Rendering via a Unified Streaming Architecture
  ACM Transactions on Architecture and Code Optimization, 2024

- Carlos Mauricio Villegas Burgos, Pei Xiong, Liangyu Qiu, Yuhao Zhu, Nick Vamivakas
  Co-designed Metaoptoelectronic Deep Learning
  Optical Express, 2023, (60)15:4356-4365

- Qiang Liu, Yuhui Hao, Weizhuang Liu, Bo Yu, Yiming Gan, Jie Tang, Shao-Shan Liu, Yuhao Zhu
  An Energy Efficient and Runtime Reconfigurable Accelerator for Robotic Localization
  IEEE Transactions on Computers, 2022, 72(7), 1943-1957.

- Shaoshan Liu, Bo Yu, Jie Tang, Yuhao Zhu, Xue Liu
  Communication Challenges in Infrastructure-Vehicle Cooperative Autonomous

Driving: A Field Deployment Perspective
IEEE Wireless Communications, 2022 May, 5;29(4):126-31.

- Carlos Mauricio Villegas Burgos, Tianqi Yang, Yuhao Zhu, Nick Vamivakas
  A Design Framework for Metasurface Optics-based Convolutional Neural Networks
  Applied Optics, 2021, (60)15:4356-4365
- Zishen Wan, Bo Yu, Thomas Yuang Li, Jie Tang, Yuhao Zhu, Yu Wang, Arijit Raychowdhury, Shaoshan Liu
  A Survey of FPGA-Based Robotic Computing
  IEEE Circuits and Systems Magazine, 2021, 21(2):48-74
- Yue Leng, Chi-chun Chen, Qiuyue Sun, Jian Huang, Yuhao Zhu
  Energy-Efficient Video Processing for Virtual Reality
  IEEE Micro Special Issue on Top Picks from the 2019 Computer Architecture Conferences, May/June 2020, 40(3):30-36
- Yiming Gan, Yuxian Qiu, Jingwen Leng, Yuhao Zhu
  SVSoC: Speculative Vision Systems-on-a-Chip
  Computer Architecture Letters (CAL), March 2019, 18(1):47-50
- Yuhao Zhu, Vijay Janapa Reddi
  Optimizing General-Purpose CPUs for Energy-Efficient Mobile Web Computing
  ACM Transactions on Computer Systems (TOCS), March 2017, 35(1):1
- Yuhao Zhu, Vijay Janapa Reddi, Robert Adolf, Saketh Rama, Brandon Reagen, Gu-Yeon Wei, David M. Brooks
  Cognitive Computing Safety: The New Horizon for Reliability / The Design and Evolution of Deep Learning Workloads
  IEEE Micro Special Issue on Cognitive Architectures, Jan/Feb 2017, 37(1):15-21
- Peter Bailis, Jean Yang, Vijay Janapa Reddi, Yuhao Zhu
  Research for Practice: Web Security and Mobile Web Computing (Website)
  Communications of the ACM (CACM), Jan 2017, 60(1):50-53
  Also appears in *ACM Queue*, July/Aug 2016, 14(4):80-95.
- Yuhao Zhu, Matthew Halpern, Vijay Janapa Reddi
  The Role of the CPU in Energy-Efficient Mobile Web Browsing
  IEEE Micro Special Issue on Mobile Systems, Jan/Feb 2015, 35(1):26-33
- Yuhao Zhu, Aditya Srikanth, Jingwen Leng, Vijay Janapa Reddi
  Exploiting Webpage Characteristics for Energy-Efficient Mobile Web Browsing
  Computer Architecture Letters (CAL), Oct 2012, 13(1):33-36
  Awarded Best of Computer Architecture Letter in 2014
- Yuhao Zhu, Bo Wang, Yangdong Deng
  Massively Parallel Logic Simulation with GPUs
  ACM Transactions on Design Automation of Electronic Systems (TODAES), June 2011, 16(3):29

Conference Papers (Co)-Led By My Group

- Ethan Chen, Jiwon Chang, Yuhao Zhu
  CoolerSpace: A Language for Physically Correct and Computationally Efficient Color Programming
  OOPSLA 2024

- Yuhao Zhu
  Invited Paper: Imaging, Computing, and Human Perception: Three Agents to Usher in the Autonomous Machine Computing Era
  ICCAD 2024
- Yuhao Zhu, Ethan Chen, Colin Hascup, Yukang Yan, Gaurav Sharma
  Computational Trichromacy Reconstruction: Empowering the Color-Vision Deficient to Recognize Colors Using Augmented Reality
  UIST 2024
- Yu Feng, Tianrui Ma, Yuhao Zhu, Xuan Zhang
  BlissCam: Boosting Eye Tracking Efficiency with Learned In-Sensor Sparse Sampling
  ISCA 2024
- Yu Feng, Zihan Liu, Jingwen Leng, Minyi Guo, Yuhao Zhu
  Cicero: Addressing Algorithmic and Architectural Bottlenecks in Neural Rendering by Radiance Warping and Memory Optimizations
  ISCA 2024
- Nisarg Ujjainkar, Ethan Shahan, Kenneth Chen, Budmonde Duinkharjav, Qi Sun, Yuhao Zhu
  Exploiting Human Color Discrimination for Memory- and Energy-Efficient Image Encoding in Virtual Reality
  ASPLOS 2024
- Abhishek Tyagi, Reiley Jeyapaul, Chuteng Zhou, Paul Whatmough, Yuhao Zhu
  Characterizing Soft-Error Resiliency in Arm's Ethos-U55 Embedded Machine Learning Accelerator
  ISPASS 2024
- Yu Feng, Tianrui Ma, Adith Boloor, Yuhao Zhu, Xuan Zhang
  Invited Paper: Learned In-Sensor Visual Computing: From Compression to Eventification
  ICCAD 2023
- Kenneth Chen, Budmonde Duinkharjav, Nisarg Ujjainkar, Ethan Shahan, Abhishek Tyagi, Jiayi He, Yuhao Zhu, Qi Sun
  Imperceptible Color Modulation for Power Saving in VR/AR
  SIGGRAPH Emerging Technologies 2023
- Tianrui Ma, Yu Feng, Xuan Zhang, Yuhao Zhu
  CamJ: Enabling System-Level Energy Modeling and Architectural Exploration for In-Sensor Visual Computing
  ISCA 2023
- Nisarg Ujjainkar, Jingwen Leng, Yuhao Zhu
  ImaGen: A General Framework for Generating Memory- and Power-Efficient Image Processing Accelerators
  ISCA 2023
- Yuhui Hao, Yiming Gan, Bo Yu, Qiang Liu, Shaoshan Liu, Yuhao Zhu
  BLITZCRANK: Factor Graph Accelerator for Motion Planning
  DAC 2023
- Yuhao Zhu
  Teaching Color Science to EECS Students Using Interactive Tutorials: Tools and

Lessons
IS&T EI (VDA) 2023

- Elias Neuman-Donihue, Michael Jarvis, Yuhao Zhu
  FastPoints: A State-of-the-Art Point Cloud Renderer for Unity
  IS&T EI (VDA) 2023
- Abhishek Tyagi, Yiming Gan, Shaoshan Liu, Bo Yu, Paul Whatmough, Yuhao Zhu
  Thales: Formulating and Estimating Architectural Vulnerability Factors for DNN Accelerators
  HPCA 2023
- Yu Feng, Patrick Hansen, Paul N. Whatmough, Guoyu Lu, Yuhao Zhu
  Fast and Accurate: Video Enhancement Using Sparse Depth
  WACV 2023
- Budmonde Duinkharjav, Kenny Chen, Abhishek Tyagi, Jiayi He, Yuhao Zhu, Qi Sun
  Color-Perception-Guided Display Power Reduction for Virtual Reality
  SIGGRAPH Asia 2022
- Yu Feng, Gunnar Hammonds, Yiming Gan, Yuhao Zhu
  Crescent: Taming Memory Irregularities for Accelerating Deep Point Cloud Analytics
  ISCA 2022
- Gregory Heyworth, Keith T. Knox, Kenneth Boydston, Yuhao Zhu
  Multispectral Scheimpflug: Imaging Degraded Books That Open Less Than 30 Degrees
  IS&T Archiving 2022
- Yiming Gan, Paul Whatmough, Bo Yu, Shaoshan Liu, Yuhao Zhu
  BRAUM: Analyzing and Protecting Autonomous Machine Software Stack
  ISSRE 2022
- Yu Feng, Nathan Goulding-Hotta, Asif Khan, Hans Reyserhove, Yuhao Zhu
  Real-Time Gaze Tracking with Event-Driven Eye Segmentation
  IEEE VR 2022
  Best Paper Nominee; Invited Presentation at IEEE VIS 2022
- Yuhao Zhu
  RTNN: Accelerating Neighbor Search Using Hardware Ray Tracing
  PPoPP 2022
- Sifan Ye, Ting Wu, Michael Jarvis, Yuhao Zhu
  Digital Reconstruction of Elmina Castle for Mobile Virtual Reality via Point-based Detail Transfer
  IS&T EI (VDA) 2022
  Kostas Pantazos Memorial Award for Outstanding Paper in Visualization and Data Analysis
- Weizhuang Liu, Bo Yu, Yiming Gan, Qiang Liu, Jie Tang, Shaoshan Liu, Yuhao Zhu
  Archytas: A Framework for Synthesizing and Dynamically Optimizing Accelerators for Robotic Localization
  MICRO 2021
- Joshua Romphf, Elias Neuman-Donihue, Gregory Heyworth, Yuhao Zhu

Resurrect3D: An Open and Customizable Platform for Visualizing and Analyzing Cultural Heritage Artifacts
Web3D 2021

- Yiming Gan, Bo Yu, Boyuan Tian, Leimeng Xu, Wei Hu, Shaoshan Liu, Qiang Liu, Yanjun Zhang, Jie Tang, Yuhao Zhu
  Eudoxus: Characterizing and Accelerating Localization in Autonomous Machines
  HPCA 2021

- Yu Feng, Boyuan Tian, Tiancheng Xu, Paul Whatmough, Yuhao Zhu
  Mesorasi: Architecture Support for Point Cloud Analytics via Delayed-Aggregation
  MICRO 2020

- Bo Yu, Wei Hu, Leimeng Xu, Jie Tang, Shaoshan Liu, Yuhao Zhu
  Building the Computing System for Autonomous Micromobility Vehicles: Design Constraints and Architectural Optimizations
  MICRO 2020

- Yiming Gan, Yuxian Qiu, Jingwen Leng, Minyi Guo, Yuhao Zhu
  Ptolemy: Architecture Support for Robust Deep Learning
  MICRO 2020

- Yu Feng, Shaoshan Liu, Yuhao Zhu
  Real-Time Spatio-Temporal LiDAR Point Cloud Compression
  IROS 2020

- Haichuan Yang, Shupeng Gui, Yuhao Zhu, Ji Liu
  Automatic Neural Network Compression by Sparsity-Quantization Joint Learning: A Constrained Optimization-based Approach
  CVPR 2020

- Yiming Gan, Yuxian Qiu, Lele Chen, Jingwen Leng, Yuhao Zhu
  Low-Latency Proactive Continuous Vision
  PACT 2020
  Best Paper Nominee

- Qiuyue Sun, Amir Taherin, Yawo Siatitse, Yuhao Zhu
  Energy-Efficient 360-Degree Video Rendering on FPGA via Algorithm-Architecture Co-Design
  FPGA 2020

- Yu Feng, Paul Whatmough, Yuhao Zhu
  ASV: Accelerated Stereo Vision System
  MICRO 2019

- Tiancheng Xu, Boyuan Tian, Yuhao Zhu
  Tigris: Architecture and Algorithms for 3D Perception in Point Clouds
  MICRO 2019
  Best Paper Nominee

- Yu Feng, Yuhao Zhu
  PES: Proactive Event Scheduling for Energy-Efficient Mobile Web Computing
  ISCA 2019

- Yue Leng, Chi-chun Chen, Qiuyue Sun, Jian Huang, Yuhao Zhu
  Energy-Efficient Video Processing for Virtual Reality
  ISCA 2019
  IEEE Micro Top Picks of Computer Architecture in 2019

- Haichuan Yang, Yuhao Zhu, Ji Liu
  ECC: Energy-Constrained Deep Neural Network Compression via a Bilinear Regression Model
  CVPR 2019
- Haichuan Yang, Yuhao Zhu, Ji Liu
  Energy-Constrained Compression for Deep Neural Networks via Weighted Sparse Projection and Layer Input Masking
  ICLR 2019
- Yue Leng, Chi-chun Chen, Qiuyue Sun, Jian Huang, Yuhao Zhu
  Semantic-Aware Virtual Reality Video Streaming
  APSys 2018
- Yuhao Zhu, Anand Samajdar, Matthew Mattina, Paul Whatmough
  Euphrates: Algorithm-SoC Co-Design for Low-Power Mobile Continuous Vision
  ISCA 2018
  IEEE Micro Top Picks of Computer Architecture (Honorable Mention) in 2018
- Yuhao Zhu, Matthew Mattina, Paul Whatmough
  Mobile Machine Learning Hardware at ARM: A Systems-on-Chip (SoC) Perspective
  SysML 2018

Conference Papers My Group Contributed To

- Shaoshan Liu, Yuhao Zhu, Bo Yu, Jean-Luc Gaudiot, Guang R. Gao
  Invited Paper: Dataflow Accelerator Architecture for Autonomous Machine Computing
  ICCAD 2024
- Yue Guan, Yuxian Qiu, Jingwen Leng, Fan Yang, Shou Yu, Yunxin Liu, Yu Feng, Yuhao Zhu, Lidong Zhou, Yun Liang, Chen Zhang, Chao Li, Minyi Guo
  Amanda: Unified Instrumentation Framework for Deep Neural Networks
  ASPLOS 2024
- Zihan Liu, Wentao Ni, Jingwen Leng, Yu Feng, Cong Guo, Quan Chen, Chao Li, Minyi Guo, Yuhao Zhu
  Juno: Optimizing High-Dimensional Approximate Nearest Neighbour Search with Sparsity-Aware Algorithm and Ray-Tracing Core Mapping
  ASPLOS 2024
- Cong Guo, Jiaming Tang, Weiming Hu, Jingwen Leng, Chen Zhang, Fan Yang, Yunxin Liu, Minyi Guo, Yuhao Zhu
  OliVe: Accelerating Large Language Models via Hardware-friendly Outlier-Victim Pair Quantization
  ISCA 2023
- Cong Guo, Chen Zhang, Jingwen Leng, Zihan Liu, Fan Yang, Yunxin Liu, Minyi Guo, Yuhao Zhu
  ANT: Exploiting Adaptive Numerical Data Type for Low-bit Deep Neural Network Quantization
  MICRO 2022
  IEEE Micro Top Picks of Computer Architecture (Honorable Mention) in 2022
- Zhi-Gang Liu, Paul Whatmough, Yuhao Zhu, Matt Mattina

S2TA: Exploiting Structured Sparsity for Energy-Efficient Mobile CNN Acceleration
HPCA 2022
Best Paper Nominee

- Yue Guan, Zhengyi Li, Zhouhan Lin, Yuhao Zhu, Jingwen Leng, Minyi Guo
  Block-Skim: Efficient Question Answering for Transformer
  AAAI 2022
- Cong Guo, Yuxian Qiu, Jingwen Leng, Xiaotian Gao, Chen Zhang, Yunxin Liu, Fan Yang, Yuhao Zhu, Minyi Guo
  SQuant: On-the-Fly Data-Free Quantization via Diagonal Hessian Approximation
  ICLR 2022
- Yangjie Zhou, Mengtian Yang, Cong Guo, Jingwen Leng, Yun Liang, Quan Chen, Minyi Guo, Yuhao Zhu
  Characterizing and Demystifying the Implicit Convolution Algorithm on Commercial Matrix-Multiplication Accelerators
  IISWC 2021
- Shaoshan Liu, Bo Yu, Yahui Liu, Kunai Zhang, Yisong Qiao, Thomas Yuang Li, Jie Tang, Yuhao Zhu
  Brief Industry Paper: The Matter of Time — A General and Efficient System for Precise Sensor Synchronization in Robotic Computing
  RTAS 2021
- Cong Guo, Bo Yang Hsueh, Jingwen Leng, Yuxian Qiu, Yue Guan, Zehuan Wang, Xiaoying Jia, Xipeng Li, Minyi Guo, Yuhao Zhu
  Accelerating Sparse DNN Models Without Hardware-Support via Tile-wise Sparsity
  SC 2020
- Cong Guo, Yangjie Zhou, Jingwen Leng, Yuhao Zhu, Zidong Du, Quan Chen, Chao Li, Minyi Guo, Bin Yao
  Balancing Efficiency and Flexibility for DNN Acceleration via Temporal GPU-Systolic Array Integration
  DAC 2020
- Anand Samajdar, J. Joseph, Yuhao Zhu, Paul Whatmough, Matt Mattina, Tushar Krishna
  A Systematic Methodology for Characterizing Scalability of DNN Accelerators
  ISPASS 2020
- Yuxian Qiu, Jingwen Leng, Cong Guo, Quan Chen, Chao Li, Minyi Guo, Yuhao Zhu
  Adversarial Defense Through Network Profiling Based Path Extraction
  CVPR 2019
- Yu Wang, Yuhao Zhu, Glenn Ko, Brandon Reagen, Gu-Yeon Wei, David Brooks
  Demystifying Bayesian Inference Workloads
  ISPASS 2019
- Yuwei Hu, Jidong Zhai, Dinghua Li, Yifan Gong, Yuhao Zhu, Wei Liu, Lei Su, Jiangming Jin
  BitFlow: Exploiting Vector Parallelism for Binary Neural Networks on CPU
  IPDPS 2018

Conference Papers From Graduate School

- Wenzhi Cui, Daniel Richins, Yuhao Zhu, Vijay Janapa Reddi
  Tail Latency in Node.js: Energy Efficient Turbo Boosting for Long Latency Requests in Event-Driven Web Services
  VEE 2019
- Yuhao Zhu, Vijay Janapa Reddi
  GreenWeb: Language Extensions for Energy-Efficient Mobile Web Computing
  PLDI 2016
- Matthew Halpern, Yuhao Zhu, Vijay Janapa Reddi
  Mobile CPU's Rise to Power: Quantifying the Impact of Generational Mobile CPU Design Trends on Performance, Energy, and User Satisfaction
  HPCA 2016
- Yuhao Zhu, Daniel Richins, Matthew Halpern, Vijay Janapa Reddi
  Microarchitectural Implications of Event-driven Server-side Web Applications
  MICRO 2015
  IEEE Micro Top Picks of Computer Architecture (Honorable Mention) in 2015
- Yuhao Zhu, Matthew Halpern, Vijay Janapa Reddi
  Event-based Scheduling for Energy-Efficient QoS (eQoS) in Mobile Web Applications
  HPCA 2015
- Matthew Halpern, Yuhao Zhu, Ramesh Peri, Vijay Janapa Reddi
  Mosaic: Cross-Platform User-Interaction Record and Replay Tool for the Fragmented Android Ecosystem
  ISPASS 2015
- Yuhao Zhu, Vijay Janapa Reddi
  WebCore: Architectural Support for Mobile Web Browsing
  ISCA 2014
- Chen Zhou, Xiaofei Wang, Weichao Xu, Yuhao Zhu, Vijay Janapa Reddi, Chris Kim
  Estimation of Instantaneous Frequency Fluctuation in a Fast DVFS Environment Using an Empirical BTI Stress-Relaxation Model
  IRPS 2014
- Yuhao Zhu, Vijay Janapa Reddi
  High-Performance and Energy-Efficient Mobile Web Browsing on Big/Little Systems
  HPCA 2013
- Yuhao Zhu, Yangdong Deng, Yubei Chen
  Hermes: An Integrated CPU/GPU Microarchitecture for IP Routing
  DAC 2011
- Bo Wang, Yuhao Zhu, Yangdong Deng
  Distributed Time, Conservative Parallel Logic Simulation on GPUs
  DAC 2010

Patents

- Yuhao Zhu, Paul Whatmough
  Region of Interest Determination in Video

US Patent App. 15/875,464

- Yuhao Zhu, Paul Whatmough
  Computer Vision Processing
  US Patent App. 16/127,007

Book Chapters

- Yangdong Deng, Yuhao Zhu, Bo Wang
  Asynchronous Parallel Logic Simulation on Modern Graphics Processors
  GPU Solutions to Multi-scale Problems in Science and Engineering, 2013
- Yangdong Deng, Xiaomemg Jiao, Shuai Mu, Kang Kang, Yuhao Zhu
  NPGPU: Network Processing on Graphics Processing Units
  Theoretical and Mathematical Foundations of Computer Science, 2011

TALKS
decks
videos

Invited Talks

- **Human Visual Perception Meets AR/VR**
  AR/VR Symposium, University of Rochester, Oct 2024, Rochester, NY
- **Perceptual Rendering Meets Neural Rendering and Architecture**
  Meta, Oct 2024, Virtual
- **Harnessing Imaging, Computing, and Biological Perception Symbiosis**
  Cornell University, Oct 2023, Ithaca, NY
  The University of Texas at Austin, Oct 2023, Austin, TX
  University of Pennsylvania, Oct 2023, Philadelphia, PA
  Duke University, Oct 2023, Durham, NC
  University of California, Los Angeles, Oct 2023, Los Angeles, CA
  University of California, Irvine, Oct 2023, Irvine, CA
  University of California, Riverside, Oct 2023, Riverside, CA
  Cambridge University, Sept 2023, Cambridge, UK
- **Harnessing and Harvesting the Computer Science-Vision Science Symbiosis**
  School of Optometry and Vision Science, UC Berkeley, Jan 2024, Berkeley, CA
  Center for Perceptual Systems, The University of Texas at Austin, Oct 2023, Austin, TX
  University of California, Santa Barbara, Oct 2023, Santa Barbara, CA
  American University, Oct 2023, Washington, D.C.
  National Eye Institute, Oct 2023, Washington, D.C.
  Dept. of Brain and Cognitive Sciences, University of Rochester, Aug 2023, Rochester, NY
  Institute of Ophthalmology, University College London, Sept 2023, London, UK
  Color Impact 2023, June 2023, Rochester, NY
  Munsell Color Science Laboratory, RIT, March 2023, Rochester, NY
  Wu Tsai Neurosciences Institute, Stanford University, April 2023, Palo Alto, CA
  Smith Kettlewell Eye Research Institute, April 2023, San Francisco, CA
- **Dark Silicon: What It Is, How We Got Here, and How We Get Out of It**
  Dept. of Physics and Astronomy, University of Rochester, Oct 2023, Rochester, NY
- **Rethinking Imaging-Computing Interface**
  ICCAD 2023, Special Session on In-Sensor AI Computing Towards Next Generation Autonomous Edge Intelligence, Nov 2023, Santa Clara, CA

Omnivision, Nov 2023, Santa Clara, CA
Stanford Center for Image Systems Engineering, April 2023, Palo Alto, CA
University College London, March 2023, Virtual

- **High Performance Embedded Imaging: An Optics, Sensing, and Computing Co-Designed Approach**
  High Performance Computing for Imaging at EI 2023, Jan 2023, San Francisco, CA
- **Color Perception-Guided Display Optimizations: Power and Beyond**
  Meta, Sept 2022, Virtual
- **Go Horizontal, Not Vertical: Addressing Visual Computing Challenges in Autonomous Machines**
  RoboArch Workshop co-located with MICRO 2022, Oct 2022, Chicago, IL
- **RTNN: Accelerating Neighbor Search Using Hardware Ray Tracing**
  WDDSA Workshop co-located with MICRO 2022, Oct 2022, Chicago, IL
- **Improving and Harnessing Software Resiliency in Autonomous Machines**
  Plenary Panel on Reliability of Autonomous Machines, IEEE COMPSAC, June 2022, Virtual
- **Visual Computing: A Horizontal Approach**
  Efficient AI Seminar, Rutgers University, Oct 2022, Virtual
  Keynote Speech, IFIP NPC, Sept 2022, Virtual
  GlobalFoundries, May 2022, Virtual
  University of Massachusetts Amherst, May 2022, Virtual
- **Intelligent Visual Computing**
  Duke University, March 2021, Virtual
  University of Utah, Oct 2021, Virtual
  HALO workshop co-located with ICCAD, Nov 2021, Virtual
- Watt-Wise Web: Architecting a Responsive and Energy-Efficient Mobile Web
  University of Utah, Oct 2020, Virtual
- Architecture Support for Robust Deep Learning: Exploiting Software 1.0 Techniques to Defend Software 2.0
  AMD Research, Oct 2020, Virtual
- **Rethinking Computer Systems Stack for Point Cloud Processing**
  Arm Research, Sept 2019, Waltham, MA
- **Getting Computer Systems Ready for Visual Computing in Ten Years**
  Intel Labs, Sept 2019, Hillsboro, OR
  Harvard University, Sept 2019, Cambridge, MA
  Yale University, Oct 2019, New Haven, CT
  University of Michigan, Nov 2019, Ann Arbor, MI
  FastPath Workshop co-located with ISPASS, Aug 2020, Virtual
- **The Next Quintillion Pixels and Beyond: Architecting Next-Generation Mobile Visual Computing Systems**
  Arm Research, Sept 2019, Austin, TX
  UT Austin, Sept 2019, Austin, TX
  Rice University, Sept 2019, Houston, TX

- **Resource-Guaranteed Deep Learning**
  Arm Research, April 2019, Waltham, MA
- **Energy-Efficient Mobile Web: Proactive and Reactive Perspectives**
  Google, April 2018, Seattle, WA
- **Algorithm-SoC Co-design for Energy-Efficient Mobile Continuous Vision**
  Cornell University, Feb 2018, Ithaca, NY
  CogArch Workshop co-located with ASPLOS 2018, March 2018, Williamsburg, VA
  SRI International, April 2018, Princeton, NJ
  Rochester Institute of Technology, Feb 2019, Rochester, NY
- **The Watt Wise Web**
  Texas A&M University, Jan 2017, Teleseminar
  Boston Area Architecture Workshop (BARC), Jan 2017, Cambridge, MA
  ARM Research, Jan 2017, Austin, TX
  UT Austin School of Information, Feb 2017, Austin, TX
- **WebCore: Architectural Support for Mobile Web Browsing**
  Intel, July 2014, Austin, TX
- **High-Performance and EnergyEfficient Mobile Web Browsing on Big/Little Systems**
  UT Austin Programming Language Lunch Seminar, September 2012, Austin, TX
  AMD Research Lab, August 2012, Austin, TX

Conference Presentations

- Computational Trichromacy Reconstruction: Empowering the Color-Vision Deficient to Recognize Colors Using Augmented Reality
  UIST 2024, Oct. 2024, Pittsburgh, PA
- Color Perception-Driven Energy Optimization for Virtual Reality: Displays and Image Encoding
  Human Vision and Electronic Imaging at EI 2024, Jan 2024, San Francisco, CA
- Modeling and Reducing Energy Consumption of Computational Image Sensors: A Case Study on Gaze Tracking
  Image Sensors and Systems at EI 2024, Jan 2024, San Francisco, CA
- Teaching Color Sciences to EECS Students Using Interactive Tutorials
  Visual Data Analytics at EI 2023, Jan 2023, San Francisco, CA
- RTNN: Accelerating Neighbor Search Using Hardware Ray Tracing
  PPoPP 2022, April 2022, Virtual
- Tail Latency in Node.js: Energy Efficient Turbo Boosting for Long Latency Requests in Event-Driven Web Services
  VEE 2019, April 2019, Providence, RI
- Algorithm-SoC Co-design for Energy-Efficient Mobile Continuous Vision
  ISCA 2018, June 2018, Los Angeles, CA
- GreenWeb: Language Extensions for Energy-Efficient Mobile Web Computing
  PLDI 2016, June 2016, Santa Barbara, CA
- Energy and Power Measurement on Mobile Devices

MobiTools co-located with ISCA 2016, June 2016, Seoul, Korea

- The Human Processing Unit (HPU) as a New Approximate Computing Substrate
  WAX 2016 co-located with ASPLOS 2016, April 2016, Atlanta, GA
- Scalable End-to-end Quality Control in Approximate Computing
  WAX 2016 co-located with ASPLOS 2016, April 2016, Atlanta, GA
- Microarchitectural Implications of Event-driven Server-side Web Applications
  MICRO 2015, December 2015, Waikiki, HI (lightening version)
- Exploiting Webpage Characteristics for Energy-Efficient Mobile Web Browsing
  Best of CAL 2014 presented at HPCA 2015, February 2015, San Francisco, CA
- Event-Based Scheduling for Energy-Efficient QoS (eQoS) in Mobile Web Applications
  HPCA 2015, February 2015, San Francisco, CA
- WebCore: Architectural Support for Mobile Web Browsing
  ISCA 2014, June 2014, Minneapolis, MN (lightening version)
- High-Performance and EnergyEfficient Mobile Web Browsing on Big/Little Systems
  HPCA 2013, February 2013, Shenzhen, China
- Hermes: An Integrated CPU/GPU Microarchitecture for IP Routing
  DAC 2011, June 2011, San Diego, CA

SOFTWARE

See the GitHub page of the Horizon Lab.

EXTERNAL SERVICE

Program Committee

- ISCA: 2025, 2024, 2023 (ERC), 2022 (ERC), 2020, 2019, 2017 (ERC)
- HPCA: 2024 (LPC), 2022, 2019, 2018 (ERC)
- MICRO: 2022 (+SRC), 2021 (ERC), 2020 (ERC), 2019
- ASPLOS: 2024
- SC: 2022
- IPDPS: 2023
- HPCI: 2022
- ECCV: 2022
- ICCV: 2023, 2021
- IEEE ISMAR: 2024
- IEEE VR: 2023
- CVPR: 2022, 2021, 2020
- ICLR: 2022, 2021, 2020
- AAAI: 2023, 2022, 2021
- NeurIPS: 2022, 2021, 2020
- WACV: 2021
- ACCV: 2020
- ISLPED: 2021
- IISWC: 2021, 2019

- ISPASS: 2022, 2019
- IEEE Micro Top Picks: 2018
- CGO-PPoPP Artifacts Evaluation: 2016, 2015
- TinyToCS Volume IV, III

Solicited Reviewer

- PPoPP (2025), ACM TECS (2024), UIST (2024), IEEE TVCG (2024, 2023), JOSA A (2023), IROS (2023), IEEE TPAMI (2023), IEEE TCAD (2023), IEEE VR (2021), IEEE TCC (2022), IEEE ISMAR (2023, 2022), ACM TOCS (2022, 2019), ASPLOS (2021), IEEE TNNLS (2021, 2020), IEEE TSUSC (2020), ACM TWEB (2020), IEEE TC (2020, 2019), IEEE CAL (2022, 2021, 2020, 2018, 2017), IEEE TMC (2017), IEEE Micro (2022, 2020, 2019, 2018, 2017), ACM TACO (2019, 2018, 2016), HCOMP (2016), IEEE ESL (2015), ACM TODAES (2011), ICS (2018), DAC (2012, 2011)

Workshop and Tutorial Organizations

- Co-organizer: *Visual Computing for Computer Architects*, tutorials co-located with ASPLOS 2023, MICRO 2023, ISCA 2024
- Workshop and Tutorial Co-Chair, MICRO 2022
- Co-chair: Artifact Evaluation Committee, IISWC 2022
- Co-organizer: *RSS2: Workshop on Robustness and Safe Software 2.0* (RSS2), co-located with ASPLOS 2022, 2021
- Co-chair: *International Workshop on Performance Analysis of Machine Learning Systems* (FastPath), co-located with ISPASS 2021
- Organizer: *Infrastructure and Methodology for SoC Performance and Power Modeling*, tutorials co-located with IISWC 2018, ASPLOS 2019, ISCA 2019
- Organizer: Cognitive Edge Computing Workshop, co-located with MICRO 2017
- Co-chair: Sensors to Cloud Architectures Workshop, co-located with HPCA 2017
- Web chair: Cognitive Edge Computing Workshop, co-located with MICRO 2016
- Organizer: MobiTools Workshop, co-located with ISCA 2016

DEPARTMENTAL SERVICE

- *PhD Admissions Committee*, 2022-2023, 2021-2022, 2020-2021, 2019-2020, 2018-2019
- *MS Admissions Committee*, 2021-2022
- *Undergraduate Curriculum Committee*, 2021-2022, 2020-2021, 2019-2020
- *Industrial Affiliates Exploration*, 2020-2021
- *Lab Committee*, 2018-2019
- *Colloquium Coordinator*, 2018-2019

OUTREACH ACTIVITIES

University of Rochester Upward Bound Math and Science (`https://www.rochester.edu/college/kearnscenter/pre-college/trio-programs.html#math-science`)

- Taught three-day workshop each summer to high-schoolers from the Vanguard Collegiate High School and Wilson High School
- Used Raspberry Pi as the platform to introduce programming to students, and incrementally built simple computer vision programs such as edge detectors.

Women in Engineering Program (WEP) (`http://www.engr.utexas.edu/wep`)

- Speaker at high school-focused summer camps at UT Austin with an emphasis on inspiring high school female about engineering (~75 high school senior women)
- Mentored two female sophomores in the Graduates Linked with Undergraduates in Engineering (GLUE) Program.

TEACHING EXPERIENCE

Instructor (University of Rochester)

- Fall 2024, CSC 259/459 *Computer Imaging, Graphics, and Human Vision*
- Fall 2024, CSC 412 *Introduction to Augmented and Virtual Reality* (co-instructor)
- Spring 2023, CSC 252/452 *Computer Organization*
- Fall 2022, CSC 292/572 *Mobile Visual Computing*
- Fall 2022, CSC 412 *Introduction to Augmented and Virtual Reality* (co-instructor)
- Spring 2022, CSC 252/452 *Computer Organization*
- Fall 2021, CSC 292/572 *Mobile Visual Computing*
- Fall 2021, CSC 412 *Introduction to Augmented and Virtual Reality* (co-instructor)
- Spring 2021, CSC 414 *Selected Topics on Augmented and Virtual Reality* (co-instructor)
- Spring 2021, CSC 252/452 *Computer Organization*
- Fall 2020, CSC 292/572 *Mobile Visual Computing*
- Fall 2020, CSC 412 *Introduction to Augmented and Virtual Reality* (co-instructor)
- Spring 2020, CSC 252/452 *Computer Organization*
- Spring 2019, CSC 252/452 *Computer Organization*
- Fall 2018, CSC 292/572 *Mobile Systems Architecture*
- Spring 2018, CSC 252/452 *Computer Organization*

Teaching Assistant (UT Austin)

- Spring 2013, *Dynamic Compilation*, with Vijay Janapa Reddi
- Spring 2011, *Computer Architecture*, with Yale N. Patt
- Fall 2010, *Introduction to Embedded Systems*, with Jonathan W. Valvano

MENTORING

University of Rochester

Current Ph.D. Students:

- Abhishek Tyagi
- Nisarg Ujjainkar
- Ethan Chen
- Weikai Lin
- Han Yan

Supervised Ph.D. Dissertations:

- Yu Feng, Ph.D. 2023. Dissertation: *Systematic Optimizations for Efficient Mobile Visual Computing*. Current position: Postdoc at UR; incoming Assistant Professor at Shanghai Jiaotong University

- Yiming Gan, Ph.D. 2023. Dissertation: *Improving Fault Tolerance of Computing Systems for Autonomous Machines*. Current position: Assistant Professor at Institute of Computing Technology, Chinese Academy of Sciences

MS Alumni:

- Raj Rajwade (M.S. 2023)
- Suumil Roy (M.S. 2023)
- Shreyan Goswami (M.S. 2022)
- Zeyi Pan (M.S. 2021)
- Hanlin Gao (M.S. 2021)
- Yi Yang (M.S. 2021)
- Ting Wu (M.S. 2020; now at Ebay)
- Tianqi Yang (M.S. 2020; now at Amazon)
- Tiancheng Xu (M.S. 2020; Ph.D. at Rice University)
- Chi-chun Chen (M.S. 2019; now at Cray Inc.)
- Boyuan Tian (M.S. 2019; Ph.D. at UIUC)

Undergraduate Alumni:

- Yongzhao Wu (Summer 2023)
- Ziyang Yuan (Summer 2023)
- Samuel Lee (Spring 2023)
- Zeyu Niu (Spring 2023)
- Ethan Shahan (Summer 2022–Spring 2023)
- Christopher Bruinsma (Spring 2022)
- Jiayi He (Spring 2022–Summer 2022; MS at CMU)
- Matan Kotler-Berkowitz (Spring 2022, Fall 2022, Spring 2023)
- Jennifer Yu (Summer 2022)
- Junhua Huang (Summer 2022)
- Meisen Hu (Summer 2022)
- Ruihan Xu (Summer 2022)
- Muhammad Qasim (Summer 2022)
- Shengyi Jia (Summer 2022; MS at John Hopkins University)
- Nikhil Khanna (Summer 2022)
- Gunnar Hammonds (Summer 2022)
- Mohamed ali manai (Summer 2022)
- Siddhant Choudhary (Summer 2022)
- Edmund Sepeku (Spring 2021)
- Kharissa King (Spring 2021)
- Elias Neuman-Donihue (Summer 2021–Spring 2022)
- Samuel Triest (Spring 2020–Spring 2021; MS/Ph.D. at CMU)
- Yawo Alphonse Siatitse (Spring 2020; MS at John Hopkins University)
- Qiuyue Sun (Summer 2018–Spring 2019; now at ByteDance)
- Sifan Ye (Spring 2019–Summer 2020; MS at Stanford University)

- Weituo Kong (Fall 2019; MS at Brown University)
- Oliver Zhang (Summer 2018; transferred to Univ. of Michigan)
- Noah Helterbrand (Summer 2018)
- Tolga Furkan Aktas (Fall 2019)
- Jessica Ervin (Fall 2019)

Previously at UT Austin

- Hannah Peeler, Undergraduate student, 2016
- Janna Tulabot, Undergraduate student, 2016
- Matthew Halpern, Ph.D. student, 2013 - 2017
- Wenzhi Cui, Ph.D. student, 2015 - 2017

DISSERTATION COMMITTEE

Ph.D. Candidate

- Carlos Mauricio Villegas Burgos (Optics, proposed in 2019)

Ph.D.

- Eslam Elmitwalli (ECE, 2024): *Applications of Randomness in Hardware Security and Combinatorial Optimization*
- Irving Barron (ECE, 2024): *Design and Applications of 2D Barcodes Using Dual Modulation*
- Yuanyuan Pan (Math, 2024): *The Damped Wave Equation and Associated Polymer*
- Uday Kumar Redd Vengalam (ECE, 2023): *Extending Ising Machines for solving Machine learning problems and LDPC codes*
- Narges Mohammadi (ECE, 2024): *Integrating Learning-based and Physical Model-based Methods for Elasticity Reconstruction*
- David Lippman (Optics, 2023): *Freeform gradient-index optical design and metrology for reduced system size and weight*
- Sayak Chakraborti (CS, 2022): *Opportunistic Resource Management: Resource Utilization in Datacenters*
- Divya Ojha (CS, 2022): *Defending Against Microarchitectural Side Channel Leaks*
- Haichuan Yang (CS, 2020): *Sparse Learning for Model Optimization*
- Daniel Nikolov (Optics, 2020): *Software and Hardware Enabling the Next-Generation Near-Eye Displays*
- Kan Xu (ECE, 2020): *Power Delivery in High Current 3-D Systems*
- Hoda Sadat Ayatollahi Tabatabaei (ECE, 2018): *Energy Balancing in Wireless Networks with MIMO Communications*

M.S.

- Junfu Zheng (Optics, 2023): *Measurement of gradient-index materials with transmission deflectometry*
- Yiwen Fan (Optics, 2020): *Numerical Calculation of Zernike Polynomials and the Sample Selection Method of NURBS Spline Generation*

B.S.

- Aayush Poudel (CS, 2023): *Compressing Ray Trajectory Mapping using Bézier Curves*
- Elias Neuman-Donihue (CS, 2022): *Fast Rendering of Massive Point Clouds*

- Benned Hedegaard (CS, 2022): *Sign-Informed Semantic Mapping for Language Interaction*
- Sifan Ye (CS, 2020): *3D Reconstruction from Colored Point Clouds with Detail Transfer*
- Samuel Triest (CS, 2020): *Unsupervised Reinforcement Learning in Environments with Strong Priors*