

Yu Feng

University of Rochester
Computer Science

yfeng28@ur.rochester.edu
www.cs.rochester.edu/u/yfeng28

RESEARCH INTERESTS

My main research interest is to explore different algorithm-architecture co-designs for faster and more energy-efficient mobile vision systems. My recent interests include in-sensor computing, multi-modality mobile applications, VR/AR visual system, etc.

EDUCATION

University of Rochester *2017 - 2023*

Ph.D., Computer Science.

Advisor: Yuhao Zhu

Carnegie Mellon University *2015 - 2017*

M.S., Material Science.

Advisor: Anthony Rollett

Tianjin Polytechnic University *2011 - 2015*

B.S., Material Science.

Advisor: Xiaoping Liang

HONORS & AWARDS

Graduate Fellowship, Carnegie Mellon University *2015 - 2017*

Second-class Scholarship, Tianjin Polytechnic University *2013 - 2015*

Thrid-class Scholarship, Tianjin Polytechnic University *2012 - 2013*

WORKING EXPERIENCE

Meta Reality Lab, Research Intern @ Meta *05/2022 - 10/2022*

- Exploited human visual system and designed an efficient passthrough pipeline for meta VR/AR products.
- Designed a new sensing scheme to reduce sensing and data communication energy consumption.
- Achieved over $6\times$ data reduction in sensing while maintaining a similar visual quality on existing VR displays.

Facebook Reality Lab, Research Intern @ Facebook *08/2020 - 12/2020*

- Worked on vision algorithm-sensor co-design in the camera sensor team.
- Designed a framework to reduce compute cost and data transmission between the image sensor and backend SoC.
- Built an end-to-end software stack and achieved $5.4\times$ speedup compared to all state-of-the-art eye tracking algorithms.

GoogleAI, Intern @ Google *06/2019 - 09/2019*

- **Obtained a return intern offer.**
- Worked on “Creative Code Prefetching” and Google Wide Profile.
- Improved Google workloads by cleverly designed “software” prefetching.

PUBLICATION Conference Articles

- Yue Guan, Yuxian Qiu, Jingwen Leng, Fan Yang, Shuo 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)

- Yu Feng, Tianrui Ma, Adith Bolor, Xuan Zhang, Yuhao Zhu
Invited Paper: Learned In-Sensor Visual Computing: From Compression to Eventification (ICCAD, 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) (*Equal contribution)
- Yu Feng, Patrick Hansen, Paul N. Whatmough, Guoyu Lu, Yuhao Zhu
Fast and Accurate: Video Enhancement Using Sparse Depth (WACV, 2023)
- Yu Feng, Gunnar Hammonds, Yiming Gan, Yuhao Zhu
Crescent: Taming Memory Irregularities for Accelerating Deep Point Cloud Analytics (ISCA, 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 Honorable Mention, Invited Paper in VIS 2022)
- Yu Feng*, Boyuan Tian*, Tiancheng Xu*, Paul Whatmough, Yuhao Zhu
Mesorasi: Architecture Support for Point Cloud Analytics via Delayed-Aggregation (MICRO, 2020) (*Equal contribution)
- Yu Feng, Shaoshan Liu, Yuhao Zhu
Real-Time Spatio-Temporal LiDAR Point Cloud Compression (IROS, 2020)
- Yu Feng, Paul Whatmough, Yuhao Zhu
ASV: Accelerated Stereo Vision System (MICRO, 2019)
- Yu Feng, Yuhao Zhu
PES: Proactive Event Scheduling for Energy-Efficient Mobile Web Computing (ISCA, 2019)

Journal Articles

- Yu Feng, Sudipto Mandal, Brian Gockel, Anthony D. Rollett
Extension of the Mechanical Threshold Stress Model to Static and Dynamic Strain Aging: Application to AA5754-O
November 2017, Volume 48, Issue 11, pp 5591–5607, MMTA

RESEARCH EXPERIENCE

Department of Computer Science, University of Rochester

In-Sensor Computing Framework

10/2022 - Now

- Surveyed the design and scaling trends of CIS, pointing out new opportunities for architectural exploration.
- Proposed the first component-level energy modeling tool for CIS and validate it against real silicon.

Improving Memory Efficiency for Point Cloud DNN

07/2021 - 03/2022

- Pinpointed the memory inefficiencies of today's point cloud DNN algorithms.
- Proposed an efficient approximation algorithm for general point-based point cloud DNNs to reduce the DRAM traffic and SRAM bank conflicts.

Continuous Real-time Mobile Eye Tracking 09/2020 - 02/2021

- Designed an eye tracking pipeline to improve mobile performance by $5\times$ while retaining the state-of-the-art accuracy.
- Proposed a lightweight and robust method to detect ROI to reduce overall computation and data communication.

Architecture Support for Point Cloud DNN 01/2020 - 05/2020

- Characterized the general compute and memory inefficiencies of today's point cloud DNN algorithms.
- Proposed an efficient framework for general point-based point cloud DNNs to reduce the algorithmic computation and improve the program parallelism.
- Introduced the necessary architecture support to further speedup the proposed framework.

Real-time Compression for LiDAR-generated Point Cloud 09/2019 - 02/2020

- Proposed an efficient real-time compression algorithms for dynamic LiDAR-generated point cloud.
- Achieved high end-to-end application accuracies and outperformed the state-of-the-art algorithms.

Algorithm and Hardware Co-design for Stereo Vision 08/2018 - 05/2019

- Proposed an energy-efficient framework for continuous stereo vision to improve real-time stereo vision throughput on mobile devices.
- Introduced a new hierarchical optimization on DNN computation based on the unique characteristics in stereo vision algorithms by partial evaluation and motion compensation.
- Coupled with the new system framework, introduced a new hardware support for stereo vision application.

Proactive Event Scheduling for Mobile Web Computing 12/2017 - 07/2018

- Proposed a new scheduling framework by proactively predicting future event in order to scale up the optimization scope.
- Introduced a new model for web event prediction by combining program analysis on web with machine learning from recent history.
- Formulated our task scheduling optimization model into a constrained problem and dynamically solve the problem using integer linear programming on-the-fly.

Department of Material Science, Carnegie Mellon University

Constitutive Modeling of Plastic Strain Aging 09/2015 - 05/2017

- Introduced a new constitutive model for dynamic and static aging for metallic materials for predicting metallic aging behaviors.
- Proposed a fast optimization routine for convergence of model fitting process.
- Applied the constitutive model to various metals such as AA-5754-O and Ti-64 serial and achieved a considerable accuracy close to experiment results.

TEACHING

Teaching Assistant

- @CMU: Engineering Mechanics I (Undergraduate)
- @CMU: Introduction to Material Science (Graduate)
- @UR: Dynamic Language and Software Development (Graduate)
- @UR: Computer Organization (Graduate, twice)