

Abstract: Delivering real-time continuous vision in an energy-efficient manner is a tall order for mobile system design. To overcome the energy-efficiency barrier, we must expand the research horizon from optimizing individual accelerators toward holistically co-designing the different components in the entire SoC.

In this talk, I will discuss one promising approach along this direction. It exploits the synergies between the imaging IP and the computer vision IP to improve the system-level energy-efficiency. Harnessing the insight that changes in pixel data between consecutive frames represents visual motion, we first propose an algorithm that leverages this motion information to relax the number of expensive CNN inferences required by continuous vision applications. We co-design a mobile SoC architecture to maximize the efficiency of the new algorithm. I will discuss the key architectural support as well as various practical SoC-design decisions. Measurement and synthesis results show that our system achieves up to 66% SoC-level energy savings with only 1% accuracy loss.

Bio: Yuhao Zhu is an Assistant Professor of Computer Science at University of Rochester. His research focuses on domain-specific architecture for emerging visual computing domains such as Augmented/Virtual Reality and continuous vision. Zhu received the Ph.D. degree from The University of Texas at Austin. He previously held visiting researcher positions at Harvard University and ARM Research. Zhu received the Honorable Mention of the 2018 ACM SIGARCH/IEEE-CS TCCA Outstanding Dissertation Award, IEEE Micro Top Picks from Computer Architecture Conferences (2015, 2018), Best of Computer Architecture Letters Award (2014), and is a recipient of the Google Faculty Research Award 2018.