Xuanteng Huang 黄軒騰

x@huangxt.cn · huangxt · huangxt.com

Computer science P.h.D student with strong backgrounds on architecture, system and software.

- Working on GPU software designs (framework/runtime/compiler) to improve the performance and throughput of general computing and machine learning workloads.
- Interned at NVIDIA/ByteDance/GSoC, familiar with CUDA/ROCm/Debian/PyTorch, etc.

Education

Sun Yat-sen University, Guangzhou, China — 2022 - 2027 (expected)
Ph.D. student in Computer Science — Advisers: Xianwei Zhang

Sun Yat-sen University, Guangzhou, China — 2018 - 2022
B.S. in Computer Science

Experiences

Google Summer of Code 2024 — May. – Aug. 2024
ROCm maintainer of Debian — Mentor: Christian Kastner
Ship and maintain open source ROCm compute stack in the official package archive of Debian/Ubuntu and their alternatives. Bridge upstream developers and end users to provide flawless experiences with AMD GPUs in Debian.

NVIDIA — Shanghai, China, Sep. – Dec. 2022
GPU Arch Intern, Deep Learning — Mentors: Ethan Yan, Vicki Wang
Involved in developing and optimizing on kernels in *cuDNN* upstream library with Fast Kernel team, GPU Compute Arch. Achieved respect **9.51**× and **12.54**× speedups for depthwise convolution kernels on Ampere and Hopper GPUs.

ByteDance — Hangzhou, China, May. – Aug. 2022
Heterogeneous Computing Intern — Mentor: Yibo Zhou
Design and implement a system-level CUDA profiling tool (proof-of-concept) based on NVIDIA CUPTI, with **0.5**× **lower overhead** than Nsight Compute with no explicit process injection.

Publications

- [DAC '24] SMILE: LLC-based Shared Memory Expansion to Improve GPU Thread Level Parallelism
  Tianyu Guo, **Xuanteng Huang**, Kan Wu, Xianwei Zhang and Nong Xiao
- [ICCD '23] KeSCo: Compiler-based Kernel Scheduling for Multi-task GPU Applications
  Zejia Lin, Zewei Mo, **Xuanteng Huang**, Xianwei Zhang and Yutong Lu
- [ECCV '20] MINI-Net: Multiple Instance Ranking Network for Video Highlight Detection
  Fa-Ting Hong, **Xuanteng Huang**, Wei-Hong Li and Wei-Shi Zheng

Skills

- Programming Languages: C, C++, Python, CUDA
- Tools: Git, CMake, LaTeX, Docker, GDB, Vim, Bash
- Artifacts: rocm-build, Debian packages

★ Honors and Awards

CCF Elite Collegiate Award [CCF 优秀大学生/领航计划]	Octorber 2022
SYSU President Scholarship [中山大学校长奖学金]	September 2022
The First Prize Student Scholarship in SYSU $\times 2$ (top 5%) [中山大学一等奖学金]	2020, 2021
Shenzhen Stock Exchange Scholarship [深交所奖学金]	September 2020
Fist Prize (Rank 1) of IndySCC'21 Student Cluster Competition	November 2021
Honorable Mention (Rank 4) of ISC'21 Student Cluster Competition	June 2021
Second Prize of ASC'20-21 Student Cluster Competition	January 2021

♡ Professional Services

- ACM EuroSys '23, Artifact Evaluation Committee
- USENIX ATC '22, Artifact Evaluation Committee
- USENIX OSDI '22, Artifact Evaluation Committee
- IEEE NAS '24, Sub-reviewer