Yangzihao Wang

Google Inc. – 1600 Amphitheatre Parkway, Mountain View, CA 94043

+1 (530) 750 9016 • yangzihao@google.com
http://yzhwang.github.io

Education

University of California, Davis
Ph.D., Computer Science, GPA: 3.84/4.0 *2011–2016*
Advisor: Prof. John D. Owens

Beihang University
M.E., Software Engineering *2008–2011*

Beihang University
B.E., Computer Science *2003–2007*

Honors and Awards

2016: Distinguished Paper, ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming

2015: Graduate Group in Computer Science Fellowship

2014: NVIDIA Graduate Fellowship Finalist

2007: Outstanding Graduates of Colleges and Universities in Beijing *top 10%*

Experience Highlights

Google Brain
Software Engineer *Jan-2017–Present*
Working on TensorFlow infrastructure and GPU performance.

Institute of Data Analysis and Visualization, UC Davis
Graduate Student Researcher *Aug-2011–Dec-2016*
Research Topics: structure of parallelism in irregular algorithms on the GPU; programming model for graph analytics on the GPU.

- Gunrock: a stable, powerful, and forward-looking substrate for GPU-based graph-centric research and development (open-sourced at: http://gunrock.github.io);
  - Led a research team of 8 graduate students;
  - Proposed the data-centric programming model for GPU graph analytics;
  - Designed the system framework and API set;
  - Implemented several critical optimizations for the system core;
  - Achieved the best performance of any programmable GPU+graph library.
- CUDA Multi-GPU sorting system based on samplesort;
  - Designed the multi-GPU cluster sorting system with MPI and CUDA;
  - Achieved comparable performance with sorting systems on CPU cluster or supercomputer.

Google
Software Engineer Intern *Fall 2015*
Worked in search infrastructure team, built a content-free URL quality score estimation service for index selection.

DARPA
Summer Research Intern *Summer 2013–2016*
Worked on expanding features and increasing the performance of Gunrock. Worked closely with multiple research groups and companies on using Gunrock for large-scale graph data analysis on real-world datasets.

AMD Research
Co-op Engineer *Summer 2012*
Ported cudaraster (a state-of-the-art software rasterizer) to OpenCL.

State Key Lab of Virtual Reality Technology and Systems, Beihang University
Research Assistant *Aug-2009–Mar-2011*
Designed and implemented a sort-first cluster rendering system. Worked on several graphics research topics such as water wave simulation, collision detection, and rendering load balancing.

External Talks

Mini-Gunrock: A Lightweight Graph Analytics Framework on the GPU
GABB'17 *May, 2017*

Gunrock: A Fast and Programmable Multi-GPU Graph Processing Library
GTC'16 *Mar, 2016*

Gunrock: A High Performance Graph Processing Library on the GPU
Oracle Labs *Apr, 2015*

High-Performance Graph Processing Programming Model on the GPU
NVIDIA Research *Jan, 2015*

High-Performance Graph Primitives on GPUs: Design and Implementation of Gunrock
GTC'14 *Mar, 2014*

Professional Skills

Proficient:: C/C++, CUDA, TensorFlow, Python, MPI, LaTeX, git, Linux development
Familiar:: OpenGL, Spark

Professional Service

Reviewer:: TC, TPDS, JDPC, PeerJ, PLDI'18
Program Committee Member:: The 1st GPUTech Workshop at ICCSA 2016
Graph Algorithms Building Blocks Workshop at IPDPS 2018

Publications

Yangzihao Wang, Yuechao Pan, Andrew Davidson, Yuduo Wu, Carl Yang, Leyuan Wang, Muhammad Osama, Chenshan Yuan, Weitang Liu, Andy T. Riffel, and John D. Owens. Gunrock: GPU graph analytics. *ACM Transactions on Parallel Computing*, 2017.

Yuechao Pan, Yangzihao Wang, Yuduo Wu, Carl Yang, and John D. Owens. Multi-GPU graph analytics. In *Proceedings of the 31st IEEE International Parallel and Distributed Processing Symposium*, IPDPS 2017, May/June 2017.

Yangzihao Wang, Sean Baxter, and John D. Owens. Mini-gunrock: A lightweight graph analytics framework on the GPU. In *Graph Algorithms Building Blocks*, GABB 2017, May 2017.

Leyuan Wang, Yangzihao Wang, Carl Yang, and John D. Owens. A comparative study on exact triangle counting algorithms on the GPU. In *Proceedings of the 1st High Performance Graph Processing Workshop*, HPGP '16, May 2016.

Yangzihao Wang, Andrew Davidson, Yuechao Pan, Yuduo Wu, Andy Riffel, and John D. Owens. Gunrock: A high-performance graph processing library on the GPU. In *Proceedings of the 21st ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, PPoPP 2016, March 2016. Distinguished Paper.

Yuduo Wu, Yangzihao Wang, Yuechao Pan, Carl Yang, and John D. Owens. Performance characterization for high-level programming models for GPU graph analytics (best paper finalist). In *IEEE International Symposium on Workload Characterization*, IISWC 2015, October 2015. Best Paper finalist.

Carl Yang, Yangzihao Wang, and John D. Owens. Fast sparse matrix and sparse vector multiplication algorithm on the GPU. In *Graph Algorithms Building Blocks*, GABB 2015, May 2015.

Afton Geil, Yangzihao Wang, and John D. Owens. WTF, GPU! Computing Twitter's who-to-follow on the GPU. In *Proceedings of the Second ACM Conference on Online Social Networks*, COSN '14, October 2014.