Simon Wang

Email: wang.c.simon@gmail.com **Website:** https://simoncwang.github.io/
Research interests: Machine Learning & AI, Computer Graphics, AR/VR, HCI

Education

University of Maryland — **College Park, MD**
M.S., Computer Science (Current GPA: 3.83/4.0) — *Expected May 2025*
B.S., Computer Science (GPA: 3.52/4.0) — *December 2023*
University Honors — *2019-2023*
Presidential Scholarship — *2019-2023*

Skills

Programming: Python, Java, HTML, Javascript (vanilla/React), C, C# (Unity), C++, SQL
ML/AI: Pytorch, HuggingFace Transformers, AutoGen, Ollama, OpenCV, OpenAI API, Langchain, Gradio
Software/Tools: GitHub, VSCode, Fusion 360, Docker, Unity, Google Suite, MS Office, MATLAB, Arduino

Work Experience

Research Assistant — June-December 2023
University of Maryland — *College Park, MD*

- Coded software tool to annotate data visualization SVGs
- Used JavaScript, HTML, and Python to develop front-end and back-end of a web page

Student Initiated Course Facilitator — January-May 2023
University of Maryland — *College Park, MD*

- Co-taught course on creating custom shaders in Three.js (CMSC398K)
- Prepared course materials and lectured about linear algebra needed for computer graphics
- Graded and gave feedback on homework and coding assignments

Software Development Engineer Intern, Amazon's Choice — May-August 2022
Amazon — *Seattle, WA*

- Developed quality assurance tools to improve Amazon's Choice recommendation system
- Used Java, Apache Spark, and AWS to push and test code on Amazon databases
- Collaborated with 20+ software engineers in fast-paced environment with daily meetings
- Presented solution ideas and final product to Amazon's Choice team and received feedback

Projects | All Projects Webpage

LLMSpatialLayout: LLM Structured Outputs to Generate Spatial Layouts from Image Descriptions - Paper Extension — January 2025
GitHub + Report
Technologies: Python, Ollama, OpenAI API

- Improvement and extension to LLM-based spatial layout generation of the paper: Grounded Text-to-Image Synthesis with Attention Refocusing
- Leveraged structured outputs through OpenAI and Ollama APIs to demonstrate simplified and reliable spatial layout generation
- Reproduced evaluations of closed and open-source models with 200+ prompts used for paper results, showing that my implementation improved format accuracy to 100% for all models (previous high 98.5%), and validity of layouts by over 3% for a small model like Llama2:13B

MMO: An Investigation of Multi-modal Multi-agent Organization and Robust Benchmarking - Course Research Project October-December 2024
Technologies: Python, PyTorch, HuggingFace, SLURM, OpenAI API GitHub Technical Report

- Individual course project for CMSC848K - Multimodal Foundation Models (Prof. Jia-Bin Huang)
- Developed a multi-agent framework using multimodal large language models (MLLMs), using OpenAI gpt-4o to coordinate open-source MLLMs through Huggingface Transformers
- Produced an improved benchmark evaluation tool to mitigate inconsistencies in current benchmarking methods to enable more robust comparison of MLLMs

Monte Carlo Renderer and Disney Principled BRDF - Course Project September-December 2024
Technologies: Python, PyTorch Technical Report

- Implemented Monte Carlo path tracing using PyTorch from course-provided skeleton code
- Utilized concepts learned in adv. computer graphics course to render with various techniques (MC integration, multiple importance sampling, neural radiosity, inverse rendering)
- Built upon path tracing code to implement the Disney Principled BRDF (bidirectional reflectance distribution function) technique, including 10+ parameters to create fine-grained controllable shading of rendered objects

Diffusion-based Generative Video Consistency - Course Research Project January-May 2024
Technologies: Python, PyTorch, SLURM, Overleaf Paper with Rebuttal

- Investigated angles to improving state-of-the-art deep learning topics in a group of 2
- Conducted extensive literature review on diffusion-based video generation and editing
- Proposed and tested a new approach improving upon and combining previous techniques such as neural layered atlases and Uni-ControlNet
- Participated in a within-course mock-conference with two rounds of anonymous peer-reviews and ultimately completed a paper that was accepted by the Professor and TAs

VR Classroom - Course Research Project Leader January-May 2024
Technologies: Unity, C#, Meta Quest III GitHub

- Ideated and proposed project investigating the potential applications of VR for education
- Led a team of 5 to develop a Unity application to run on the Meta Quest III over the course of a semester by delegating tasks and collaborating with teammates
- Conducted an IRB-approved user study of 30+ participants, presented findings to class and wrote a technical report summarizing the research process and impacts

Publications

1. Chen Chen, Hannah K. Bako, Peihong Yu, John Hooker, Jeffrey Joyal, Simon C. Wang, Samuel Kim, Jessica Wu, Aoxue Ding, Lara Sandeep, Alex Chen, Chayanika Sinha, Zhicheng Liu. "VisAnatomy: An SVG Chart Corpus with Fine-Grained Semantic Labels." *arXiv preprint arXiv:2410.12268 (2024)*

Relevant Coursework

Multimodal Foundation Models, Deep Learning, Database System Architecture & Implementation, Human-Computer Interaction, XR, Advanced Computer Graphics, Game Programming, Data Visualization, Advanced Algorithms, Data Structures, Applied Probability & Statistics, Linear Algebra, Calculus 3