

Jerry Ma

mlfjerry2000@gmail.com | (858) 220-2622 | Portfolio | Github/Dramalf | LinkedIn/lingfeng-ma

Education

University of California, San Diego

09/2024 - 12/2025

M.S · Electrical & Computer Engineering, Machine Learning and Data Science

San Diego, CA

Coursework: Parallel Computation, Advanced Image Synthesis, Recommender Sys&Web Mining, etc.

Huazhong University of Science and Technology

09/2018 - 06/2022

B.E · Computer Science

China

Work Experience

Tiktok | Software Engineer Intern @Intelligent Creation

06/2025 - 12/2025 | San Jose, CA

- Engineered an AI Design Agent with Model Context Protocol (**MCP**) integration for design-in-the-loop on a Cocos2D-x-based editor, powered by Claude 3.7 (schedule agent) and **on-device** Gemini Nano (actor agent).
- Implemented context engineering with retrieval-augmented generation (**RAG**), enabling multi-entity context management for design scenario.
- Built a 30-task benchmark platform to evaluate MCP and design agent performance by measuring visual style consistency, dialogue turns, token usage, and creator experience.

Alibaba | Software Engineer @AI Pilot

07/2022 - 07/2024 | China

- Devised a Multimedia & AI Workflow Platform with Kubernetes and React for orchestrating and sharing consumable AI task pipelines, ranking **1st/27** in Alibaba Innovation Competition.
- Integrated AIGC tools into the in-browser graphic editor, cutting design & verification time by **40%**.
- Architected a cross-platform video composition framework with FFmpeg and Node.js for both CPU/GPU machines, adopted by **3 engineering teams**, supporting the generation of **2M+** videos and posters.
- Refactored the image-processing module in Rust & WebAssembly (WASM) with multithreading, fixing the memory leak and cutting latency **from 350 ms to 200 ms**.
- Enhanced the C++ multimedia service with heartbeat detection via RocketMQ and Redis, reducing the crash rates from **3% to 0.5%**.

Alibaba | Software Engineer Intern @Video Streaming

06/2021 - 08/2021 | China

- Developed a video editor SDK in TypeScript using FFmpeg.wasm & WebCodecs, enabling in-browser frame stream encoding. (Granted national patent)
- Accelerated the video compositing speed **3×** via async programming and WebGL.

Projects

GPU Matrix Multiplication | Performance-Optimized SGEMM in C++ & CUDA

09/2024 - 12/2024

- Implemented a CUDA SGEMM kernel with 2D block tiling and outer-product computation, achieving 4,780 GFLOPs (**62%** of NVIDIA T4 peak) and **85×** speedup over the baseline.
- Optimized global/shared memory usage and register allocation (80 registers/thread) to sustain **80%** SM occupancy for both square non-square matrices.

Physically Based Rendering | BSDF & Volumetric Path Tracing in C++

01/2025 - 03/2025

- Enhanced Disney Principled BSDF with multi-layer materials, enforcing energy conservation via Fresnel-Schlick approximation, and applying importance sampling to reduce noise and improve convergence.
- Extended Volumetric Path Tracing with multiple scattering and absorption integrating Henyey–Greenstein phase functions to render realistic flame, ash, and smoke effects from Blender scenes.

Skills

Languages: JavaScript/TypeScript (React), Rust (WebAssembly), C++, Python, Node.js, Java, Shader Language (GLSL/WSGL), HTML/CSS, SQL

Others: Git, Unix, Docker, FFmpeg, MySQL, MongoDB, RESTful API, AWS, CI/CD, Web Scraper, Postman, Browser Extension, Figma