

GPU-Based Batch Analysis of 3D Facial Scans

Author: Ing. Marek Horský

Supervisor: doc. RNDr. Radek Ošlejšek, Ph.D.

Masaryk University
Faculty of Informatics

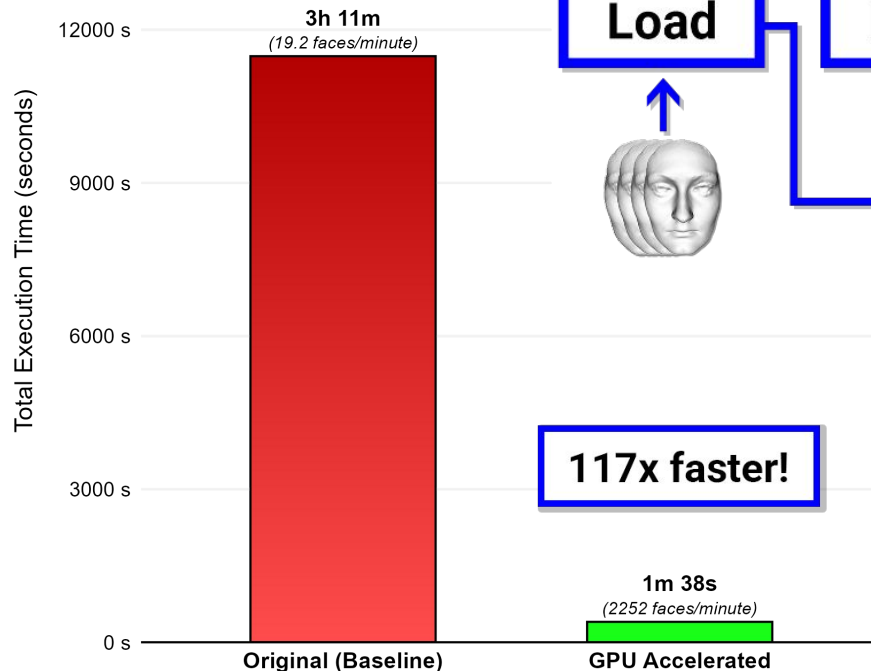


Introduction

Automated 3D face analysis experiences growing demand for high-throughput processing. FIDENTIS Analyst II is specialized software developed at the Faculty of Informatics, Masaryk University, used in anthropological research and forensic investigations. However, analyzing large datasets traditionally required hours of compute time.

This thesis reworks multiple underlying algorithms to reduce execution time from hours to mere seconds.

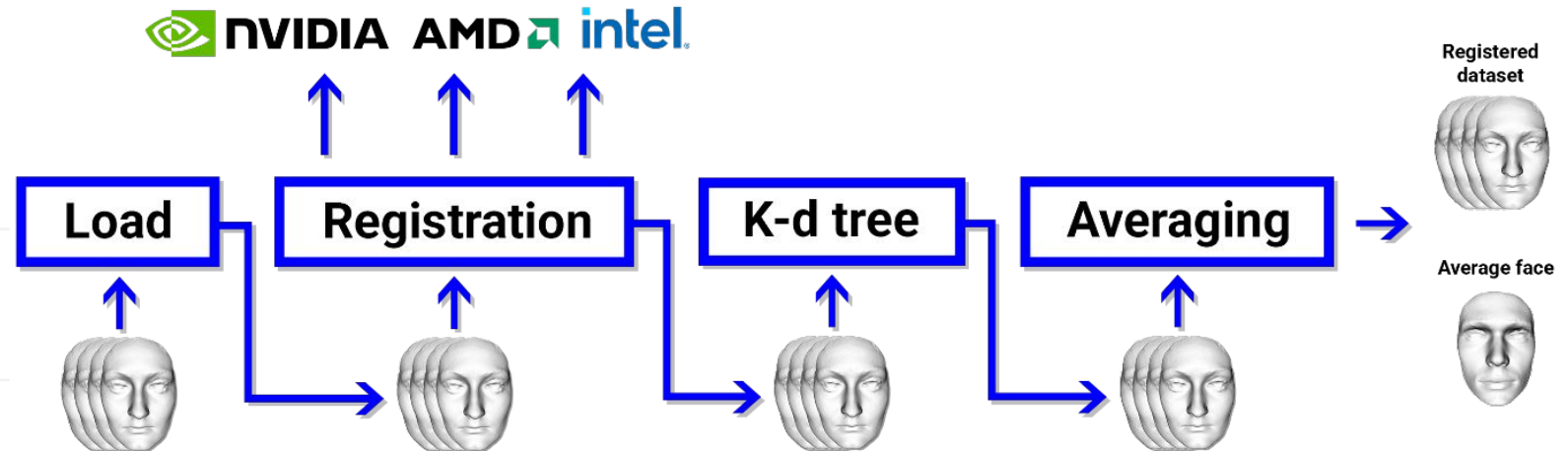
Benchmark



Solution

This thesis presents the systematic acceleration of a bottlenecked pipeline through a wide range of optimizations, spanning high-level architectural redesign down to low-level GPU kernel implementation. Key contributions include:

- **Multithreaded Java Pipeline:** Created a parallel execution model that divides and distributes analysis tasks across available threads.
- **GPU-Accelerated ICP Registration:** Implemented the Iterative Closest Point algorithm using high-performance OpenCL kernels.
- **Algorithmic Offloading:** Accelerated face averaging strategy with Nearest-Neighbor search performed on the GPU.
- **In-Place K-d Tree Construction:** Sped up spatial tree construction by optimizing in-place partitioning with the QuickSelect algorithm.
- **Data Transfer & Memory Optimization:** Applied data alignment and serialization techniques to streamline communication between SQL, Java, and GPU VRAM.
- **Multi-GPU Auto-Scaling:** Added multi-GPU scalability to FIDENTIS Analyst II, allowing seamless deployment on high-performance workstations.



Results

The presented work, evaluated on a real-world dataset of 1,226 high-resolution face scans (over 66 million 3D vertices per batch iteration), showcased a dramatic performance increase, achieving up to 117x faster batch face registration on the very same consumer-grade laptop.

Further breakdown revealed a 25x speedup in GPU-based ICP registration, a 10x speedup in face averaging, and a 10x memory footprint reduction in K-d tree construction. This was achieved through heavy parallelization, data transfer optimizations, and offloading workloads to both integrated and dedicated graphics cards, drastically accelerating anthropological and forensic workflows.