PACK It In, PACK It Out: An Experimental File Compression Method

Cestone, Joseph (School: Ridgeview High School)
Dhesi, Kevindeep (School: Ridgeview High School)

In 2018, there were 33 trillion Gigabytes of data in the world. Compression algorithms are used to store this data more efficiently. We developed and tested a new lossless compression algorithm, PACK, the predictive asynchronous compressed keyspace. We used uncompressed TIFF images exported from raw for minimal processing, simplicity and predictability for this stage of testing. The algorithm was tested against ZIP, the most common compression method, which has been updated since its development in 1989. This year, we implemented multi-threading to allow portions of file data (chunks) to be compressed simultaneously. It was hypothesized that by reducing the mathematical base of pixels, image files could be compressed smaller than with ZIP compression, and that this process could be expedited through multithreaded computing. The PACK algorithm compressed 64% of the images smaller than with the ZIP algorithm, and the multithreaded PACK version was 2.4 times faster and compressed 22% of the images smaller than ZIP. The hypothesis was correct; the PACK algorithm can often compress images smaller than the market-dominant ZIP compression algorithm, and our PACK algorithm can run faster using multithreaded computing, though it is still slower than ZIP. We believe with further improvement, our PACK algorithm can be both more effective and faster than ZIP.

Awards Won: