Parallel Robust Principal Component Analysis

Michael J. Anderson, Stephan J. Ritter, James Demmel

Berkeley Parlab

GROUP IN **BIOSTATISTICS**

Robust PCA

- Decompose a matrix into a **low rank** component and a **sparse** component:

$$M = L + S$$

= +
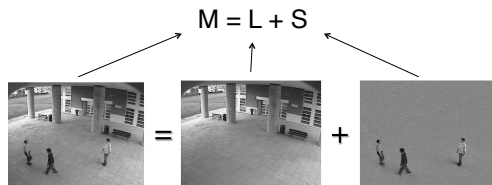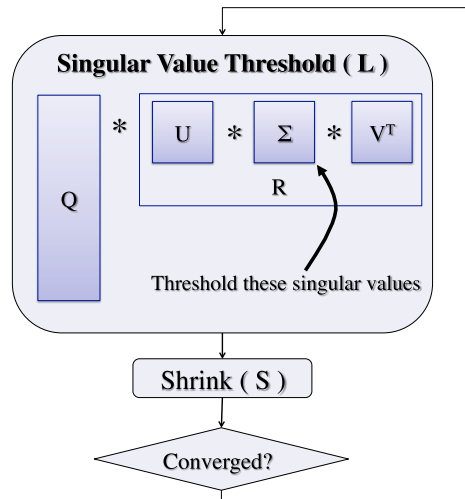
- M, L and S are **videos**
- Each column is a frame
- Dimension roughly 100,000 x N for N frames

Emmanuel Candes, 2009

Algorithm

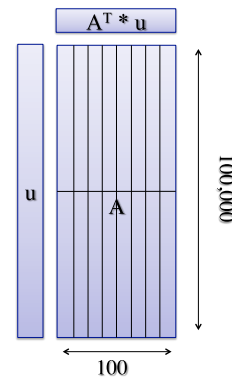minimize $\|L\|_* + \lambda\|S\|_1$
subject to $L + S = M$

Singular Value Threshold (L)

Q * [U * Σ * V^T]
R

Threshold these singular values

Shrink (S)

Converged?

- **Challenge: QR decomposition of a tall-skinny matrix**

GPU QR

- Householder QR

- for (i = 1..n)
  Form u from A$^{(i)}$
  SGEMV
  SGER
  Global Barrier

- Had to make custom SGEMV for tall-skinny matrix

A^T * u

u   A

100,000

100

Implementation	Time	Speedup
LAPACK QR (SGEQRF)	0.71 sec.	-
GPU QR	.072 sec.	**9.9x**

Matrix Dimensions Determine the Algorithms

- Matrix is generally not wide enough to provide enough parallelism in the 'n' direction. Must also divide in the 'm' direction and use atomic add when necessary.

- Best results came from minimizing the number of parallel reductions in our SGEMV

- Global barrier could potentially be avoided, but synchronization costs were a second-order concern for this size matrix

SV Threshold Optimization

Q * U * Σ * VT

Mostly set to 0

So we don't need all of this
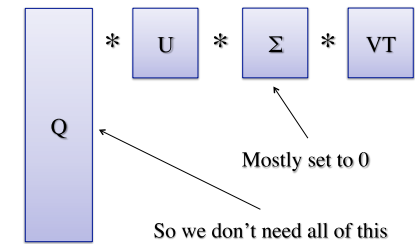
Implementation	Time	Speedup
LAPACK without optimization	1.54 sec	-
LAPACK with optimization	0.81 sec	**1.9x**
GPU with optimization	0.093 sec	**17x**

Robust PCA Speedup

Can process 10 seconds of video in 11 seconds

Went from 28x **slower** than real-time to near real-time