Advances in Video Processing using Parallelism

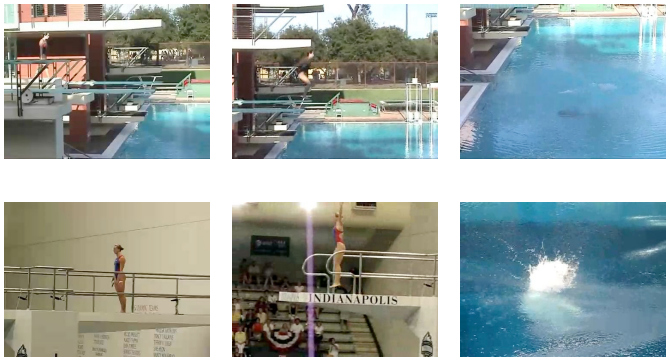
Narayanan Sundaram, Kurt Keutzer

Berkeley Parlab

PARLAS

Motion Analysis

- Motion analysis is key to semantic video understanding
- Long range motion analysis requires us to track points densely over many frames accurately
- Runtime performance is important as data grows very quickly (1080p at 30 frames/sec = 11 GB of uncompressed data per minute)
- Key-frames-based analysis reaches its limits very quickly e.g. springboard diving vs platform diving classification is very hard to do using only appearance

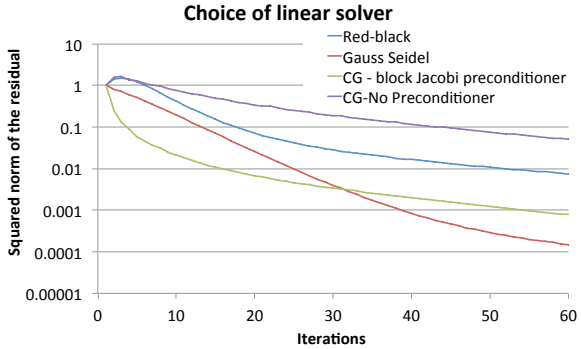


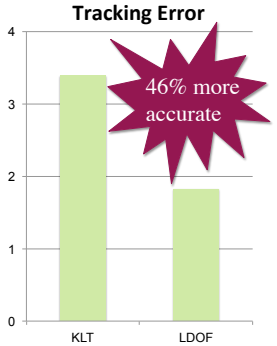
Springboard diving

Platform diving

Large Displacement Optical Flow

- Optical Flow involves computing the motion vectors ("flow field") between the consecutive frames of a video.
- We use the *Large Displacement Optical Flow (LDOF)* algorithm, which is crucial for point tracking in real world videos.
- We have implemented an efficient parallel version (in CUDA) that computes high quality flow vectors in < 2 seconds/frame
  - Possible through algorithmic exploration, numerical analysis and faster hardware
- Point tracker based on LDOF outperforms other trackers (46-66% better than other trackers)*





* Based on the MIT dataset (Liu et al, CVPR 2008)
* Based on particle trajectories from http://rvsn.csail.mit.edu/pv/data/pv/ (Sand and Teller, IJCV 2008)

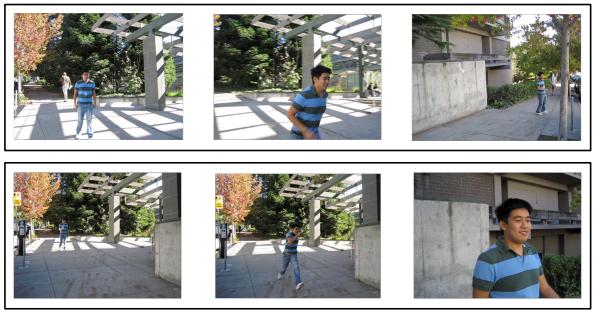
Video Object Segmentation

- Extracting video objects is crucial for video editing and semantic understanding
- We are implementing a system that combines both image contour detection and optical flow to perform video object segmentation
  - The core of the system is a sparse eigen solver
  - The resulting sparse matrix is extremely large (~30 GB for 100 frames) to fit in one GPU
- We are currently using NERSC's Dirac cluster to run the application
- Preliminary results are promising

An example sequence from "Miss Marple" and its expected output

Aesthetics analysis

- Aesthetic appeal is an important quality for human appreciation of media
- Motion is a significant metric for measuring aesthetic appeal in consumer generated videos
- We are currently generating a dataset for this purpose
- This work is being done in collaboration with Computer Graphics researchers at UC Berkeley



Two shots of a similar scene from different angles

Sports Categorization

- Determine how useful motion is for classifying sports sequences
- We used the Olympic dataset from Stanford
  - Wide variety of motions
  - Real world videos
  - Occlusions, pose changes, camera motion, background variance, compression artifacts etc.

Optimized Optical Flow → Subject Tracking → Motion histograms → Classifier

- Motion-only classifier performed just as well as an appearance based classifier (accuracy of ~33%)
- Combination of both techniques along with human detection could give even better performance