# Video Segmentation

Narayanan Sundaram, Kurt Keutzer

Berkeley Parlab

## Motivation

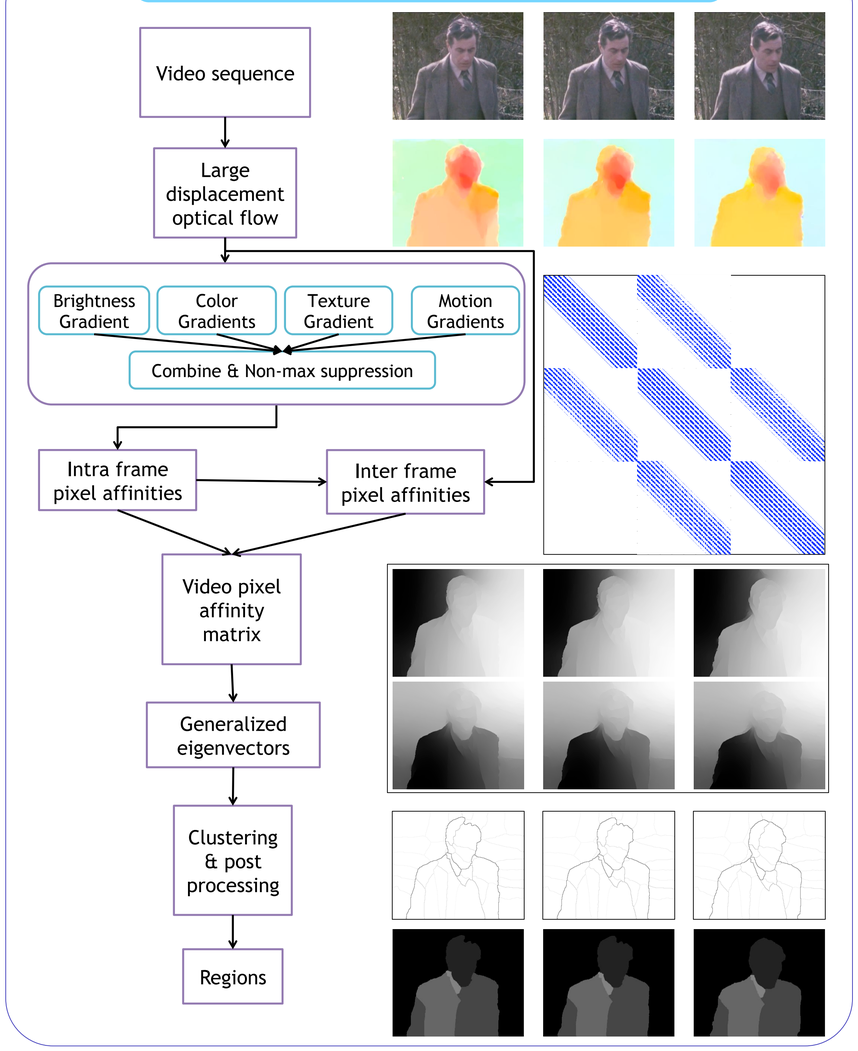
- The objective of video segmentation is to separate out objects from a video sequence based on appearance and motion
- Segmentation is required for video editing and video understanding. It is a necessary step in many applications, but computationally intensive
- We are not just parallelizing existing algorithms, but are also developing new ones
- Parallelization has enabled us to move highly accurate computer vision algorithms from images to videos
- We are using a cluster today, but expect it to run on the desktop within 5-10 years

## Example

Example video sequence

Ground truth segmentation

## Application Overview




## Computational Challenges

- Most of the computational challenges are in the eigensolver
- Large memory requirements
  - The video affinity matrix takes up ~20 GB for 100 frames of size 640x480
  - Single node solution is impractical
  - Solution: Distributed computing – Dirac cluster at NERSC
- Computing $A^Tx$ is much slower than $Ax$ for matrix blocks
  - $A^Tx$ runs about 3x slower than $Ax$
- Bugs caused due to floating point non-determinism
  - MPI_Reduce, AtomicAdd result in errors due to non-determinism

## Algorithmic Challenges

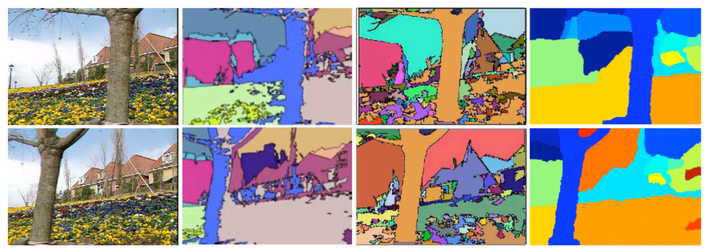
- Defining inter-frame pixel affinities
  - Combine intra-frame affinities and optical flow
- "Leakage" in eigenvectors
  - Solved through k-means clustering and Ultrametric contour maps

## Results

| Algorithm | Label density | Overall accuracy (% pixels correctly labeled) | Objects extracted with <10% error | Over-segmentation |
|---|---|---|---|---|
| Brox & Malik - trajectory segmentation [1] | 3.43% | 3.17% | 7 | 3.14 |
| ALC with incomplete tracks [2] | 3.43% | 2.77% | 0 | 54.57 |
| Hierarchical Graph Segmentation [3] | 100% | 79.23% | 0 | 10.42 |
| Our technique - Pixel level spectral clustering | 100% | 84.54% | 3 | 5.86 |

[1] T. Brox and J. Malik. Object segmentation by long term analysis of point trajectories. In ECCV, 2010.
[2] S. Rao, R. Tron, R. Vidal, and Y. Ma. Motion segmentation via robust subspace separation in the presence of outlying, incomplete or corrupted trajectories. In CVPR, 2008.
[3] M. Grundmann, V. Kwatra, M. Han, and I. Essa. Efficient hierarchical graph-based video segmentation. In CVPR, 2010.
[4] W. Brendel and S. Todorovic. Video object segmentation by tracking regions. In ICCV, 2009.

- ❑ Spectral clustering at the pixel level is practical and better than existing algorithms
  - ❑ 30x more dense than sparse clustering
  - ❑ 2x less over-segmentation than dense techniques
- ❑ Runtime is ~5 minutes for a 200 frame sequence (1.5 sec/ frame) on a 34-node GPU+multicore CPU cluster at NERSC



Original sequence | Results from [3] | Results from [4] | Our results

### Results on marple1 sequence