Action Sequence Images

Adapted From “Automatic Generation of Action Sequence Images from Burst Shots” by Sean Chen, Ben Stabler, and Andrew Stanley (Stanford University)

John Kelly
Where I Want To Be
Where I Am At (Best Case)
What’s the Goal Here?

● Input: A series of images (burst, video, etc.)
● Output: One image representing action
  o Seamless background
  o Multiple moments of the main subject(s)
  o Minimal noise around subject(s)
● Constraints:
  o Known subject count
  o Very limited camera movement
Relevant Techniques

- Image Transformations
- Color Spaces
- Thresholding
- Region Descriptors
- Morphological Processing
- Edge Detection
- Image Blending
Previous Technique

1. Align each frame
2. Extract a background
   a. (This is pretty involved…)
3. Find frame-wise differences to background
4. Compile clipped subjects onto background

Previous paper is very well done -- I am reimplementing the existing technique as given.
1. Align Each Frame

- Identify key features in each frame with SIFT
1. Align Each Frame

- RANSAC: RANdom SAmple Consensus
  - With some statistics, cleverly guess pairs of points
  - Converges quickly: ~15 iterations, ~3 seconds

Matching SIFT features are shown.
1. Align Each Frame

- Now know how to align Frame 1, Frame N
2. Extract a Background

- RGB -> HSV -> Difference* -> NormDiff
  - *Need to correct for circular hue channel
2. Extract a Background

- Threshold the normalized differences
- Apply morphological operations to clean up
2. Extract a Background

- Run edge detection on frames
- Extract silhouettes of largest objects
2. Extract a Background

- Compare silhouettes to frame edges
- Better correlation = ignore that frame data
2. Extract a Background

- Clip and Stitch/Blend
2. Extract a Background
3. Find Frame-Wise Diffs

- Very similar to finding the background
- Difference mask selects moving object
4. Compile Clipped Subjects
Progress: Okayish
Failures: Lack of Contrast
Failures: Change of Focus
Failures: Target Too Slow
Failures: Poor Threshold
Previous Work: Concerns
Difficulties On My End

- Poor Data
  - Artifacting, poor subjects, uncontrolled
- Lots of parameters
  - Thresholding, cluster counts, blending, morphological dilating/eroding/opening/closing...
- Slow Iteration
  - Working with large images -> Large processing time
Continuing Work

- Tweak the parameters to improve clipping
  - Paper has some suggestions
- Acquire controlled data
  - Few moving objects, high contrast, consistent lighting, consistent focus …
Questions