Topics to be Covered in the Rest of the Semester

A brief overview to help you think about project topics
- Structure-from-motion and photo tourism
- Stereo
- Motion and tracking
- Segmentation
- Active contours
- Markov random fields, global optimization and graph cuts
- Face recognition and detection
- Instance recognition
- Category recognition
- High-dynamic range imaging
- Image matting and compositing
- Texture analysis, synthesis, inpainting

What to Think About

As we discuss the topics in this overview, think about

- Applications: fun, interesting, useful things you would like to do with images and video
- How any one of the problems we discuss and associated techniques may be mapped to your application idea(s)
- What problems and challenges might you be able to investigate?

Structure From Motion

Problem overview
- Given either:
  - An image sequence taken from a moving camera
  - A set of images of the same scene taken from different cameras and viewpoints
- Estimate:
  - Determine which images show overlapping views
  - Estimate camera positions and internal parameters
  - Estimate 3d locations of a set of keypoints
Structure From Motion

Techniques
- Keypoint matching: image-to-image or one image against many
- Fundamental matrix estimation for each pair of matching images
- Camera estimation based on assumptions about intrinsic parameters
- Multi-image “bundle adjustment” to place points and cameras in the world.

Applications
- Photo tourism
- Match-move scene insertion

Two-Camera Stereo Correspondence

Problem
- Given: images, $I_1$ and $I_2$, taken (more or less) simultaneously from two different positions with (usually) calibrated cameras
- Problem: for each pixel in $I_1$ find the corresponding pixel in $I_2$, and for each pixel in $I_2$ find the corresponding pixel in $I_1$.
  - This is known as “disparity computation”

http://vision.middlebury.edu/stereo/data/scenes2003/

Technique
- Pixel similarity measures
- Surface smoothness, with allowances for discontinuities
- Graph-based global optimization algorithms
Two-Camera Stereo Correspondence

Applications
- View synthesis and image-based rendering
- 3d reconstruction
- Robot navigation

Motion

Problem
- Given: two or more images of a scene, where both the camera and objects in the scene may be moving
- Compute the apparent motion of the image:
  - Pixel-level offset vectors, also known as optical flow.
  - Parametric model, similar to image registration
  - Layered motion: segment image into layers of different motion, each having their own model

Techniques
- Image-to-image correlation
- Spatial and temporal image derivatives to generate image constraints
- Parametric model fitting or motion-field smoothness measures; followed by optimization
- Iterate:
  - Assign pixels to layers
  - Estimate motion within each layer

Figure 8.1 of the Szeliski text.
Motion — Applications

- View stabilization
- Removal of motion blur
- Super-resolution
- Separation of transparent motion
- Foreground-background segmentation

Tracking

**Problem**

- Given a video sequence
- Identify and continually locate moving objects in the scene
  - Could include estimating the parameters of a dynamic model of the scene

**Techniques**

- Identify and follow “interesting” features; cluster those with coherent motions
- Extract and follow contours
- Correlation
- Match distributions
- Kalman filtering

**Applications**

- Security
- Human-computer interaction
- Traffic safety
- Activity recognition
- Multiframe segmentation
Segmentation

- Separate an image into foreground region (region of interest) and a background region.
- Wide variety of techniques proposed:
  - Classical methods: thresholding, merging and splitting, watershed
  - Contour driven
  - Statistical methods and mean-shift
  - Graph-based algorithms
- Many of these can be applied as part of interactive tools / user-assist tools.

“Classical” methods

- Thresholding: single and multiple threshold
- Regional split and merge
- Watershed
- Applications of these are mostly in medical imaging

Contour Driven: Active Contours

- Compute the location of a contour $C(t)$ in an image
- Constraints are combined using what can be thought of as a “soft” form of Lagrange multipliers
  - Prefer smoother curves — low curvature preferred
  - Prefer to stay in high gradient regions
- Requires initialization
- Level-set methods to avoid topological issues
- Applications in medical imaging and in interactive segmentation methods

Segmentation — Statistical methods and mean-shift

- Form measurement vector on each pixel in a high dimensional space.
  - Location, color, texture, etc.
- Think of the distribution of these points
- Each pixel attempts to seek the mode of its distribution
- Segment based on who belongs to which mode
- Often used as a pre-processing step to “oversegment” image and create “super pixels”.
Segmentation — Graph-Based Methods

- Image forms weighted graph
- Pixels connected to neighbors, with weights based on consistency between them — color, texture, etc.
- Each pixel also connected to a special “source” (foreground) and a “sink” (background) node in the graph
- Compute the minimal “cut” in the graph that separates the source and the sink.
- Pixels connected to the source node are foreground and pixels connected to the sink are background.

Recognition

Types of recognition problems

- Face recognition: given an image (or subimage) that is known to show a face, whose face is it?
- Face detection: find all of the faces in the image
- Instance recognition: given a set of particular objects, which (if any) of them appear in the image?
  - Examples range from recognizing music CD covers to particular locations on a street
- Category recognition: find all of the cars, chairs, horses, cows, bicycles etc. in a scene.
- Context recognition: divide up an image into the streets, cars, buildings, grass, sky, etc.

Algorithmic Tools for Recognition

Face recognition

- Write image as a vector and normalize
- Gather many different vectors for different faces
- Apply principal component analysis (PCA) to the vectors
- Match by projecting new images onto the principal components

Face detection

- Learned appearance models for parts of faces
  - PCA
  - Trained classifiers using large numbers of features
  - Neural networks
- Spatial relationships among parts
Algorithmic Tools for Recognition

Instance recognition — among many different objects
- Keypoint detection and description
- Hierarchical (tree) clustering of descriptors — following from "vocabulary trees" — with leaf nodes have the (potential) instances
- Match an image by matching its keypoints against the tree
- Test those "instances" that have a lot of keypoint matches based on, for example, fundamental matrix estimation

Category recognition — a much harder problem
- Bag of words: match enough SIFT (or more recent) keypoints and descriptors and you have recognized the object, independent of the relative positions of the keypoints
- Parts models: tree-structured, graphical representation of the relationship between body parts and how they move.
- Segmentation and labeling: integrate the labeling of objects and the segmentation of the object from the background through combinations of classifiers and graphical models.

Context recognition
- Divide up an image into the streets, cars, buildings, grass, sky, etc.
- Simultaneous recognition and segmentation
- Conditional random fields — a generalization of Markov random fields
- Probability that an image region is a car depends on probability that nearby region is a street.
Context Recognition — Some Successes

Computational Photography

Ch. 10 of Szeliski, but we will only cover part
- Photometric calibration
- High dynamic range imaging
- Super-resolution and blur removal
- Image matting and compositing
- Texture analysis and synthesis

We will only cover the latter during the last part of the semester, but the other topics may be of interest for projects.

Photometric calibration
- Response curve: radiance to intensity; color balancing.
- Vignetting correction
- Spatial blur estimation

High dynamic range imaging
- Can't represent a wide enough set of radiance values in $[0, ..., 255]$.
- Take images under different exposures
- Invert response curve to get radiance
- Remap ("tone mapping") to range 0..255
Super-resolution and blur removal
- Deconvolution for single images
- Multiple images:
  - Highly-accurate, sub-pixel alignment
  - Interpolate pixel values

Image matting and compositing
- Simple view: segment region from one image, paste it on another
- Blue screen techniques have been used for years.
- Multiview and motion-based techniques are of continuing interest
- "Mixed pixels", combining foreground and background, are of ongoing difficulty for both segmentation and compositing.
  - Effectively computing an alpha channel

Natural Image Matting Example

Texture analysis and synthesis
- Given a “texture”, generate new textures that look similar
- Matching of frequency characteristics at multiple scales
- Hole-filling and inpainting
- Non-photo-realistic rendering.