395T Visual Recognition: Outline of lecture for Sept 28, 2012

I. Generic object categorization
   a. Window-based models
      i. Person detection with SVM and HOG (Dalal & Triggs, 2005)
         1. Support vector machines
         2. HOG descriptor
      ii. Pros and cons of window-based models

   b. Part-based models
      i. Bag-of-words
         1. e.g., with Naïve Bayes classifier
         2. Local feature sampling strategies for categorization
         3. Pyramid match kernel
      ii. Generalized Hough for category detection
         1. Implicit shape model (Leibe et al. 2004)
         2. (Class-specific Hough forests – Lempitsky et al.)
      iii. (Deformable part-based model with latent SVM (Felzenszwalb et al. 2008))

II. Mid-level representations
   a. Edge detection
      i. Canny example
   b. Texture representation
      i. Filter banks
      ii. Textons
   c. Segmentation into regions
      i. Gestalt properties
      ii. Segmentation as clustering, grouping
   d. Ongoing topics in mid-level visual representations

Reminder: Assignment 2 due Oct 5.

Plan for today

- Wrap-up on window- and part-based models
- Introduction to mid-level representations
- Student presentations and paper discussion

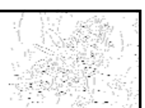
- HW1 returned

Mid-level cues

Tokens beyond pixels and filter responses but before object/scene categories

- Edges, contours
- Texture
- Regions
- Surfaces

Gradients -> edges



Primary edge detection steps:

1. Smoothing: suppress noise
2. Edge enhancement: filter for contrast
3. Edge localization

   Determine which local maxima from filter output are actually edges vs. noise

   - Threshold, Thin

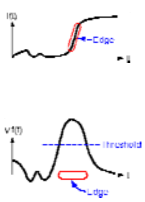
Kristen Grauman

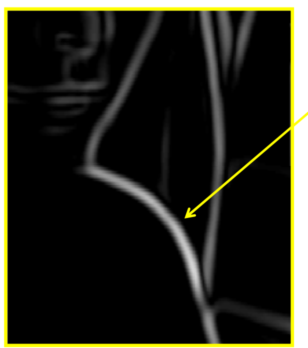
Canny edge detector

- Filter image with derivative of Gaussian
- Find magnitude and orientation of gradient
- **Non-maximum suppression**:
  - Thin wide “ridges” down to single pixel width
- **Linking and thresholding** (**hysteresis**):
  - Define two thresholds: low and high
  - Use the high threshold to start edge curves and the low threshold to continue them
- MATLAB: **`edge(image, 'canny');`**
- **`>>help edge`**

Source: D. Lowe, L. Fei-Fei

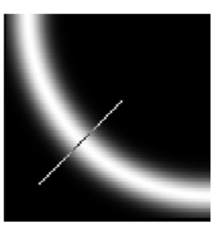
The Canny edge detector

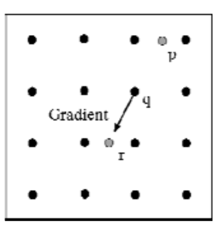




How to turn these thick regions of the gradient into curves?

Non-maximum suppression





Check if pixel is local maximum along gradient direction, select single max across width of the edge
- requires checking interpolated pixels p and r

The Canny edge detector

Problem: pixels along this edge didn't survive the thresholding

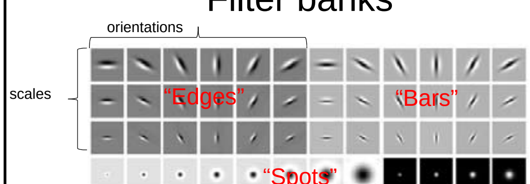
thinning
(non-maximum suppression)

Texture representation

- Textures are made up of repeated local patterns, so:
  - Find the patterns
    - Use filters that look like patterns (spots, bars, raw patches…)
    - Consider magnitude of response
  - Describe their statistics within each local window
    - Mean, standard deviation
    - Histogram
    - Histogram of "prototypical" feature occurrences

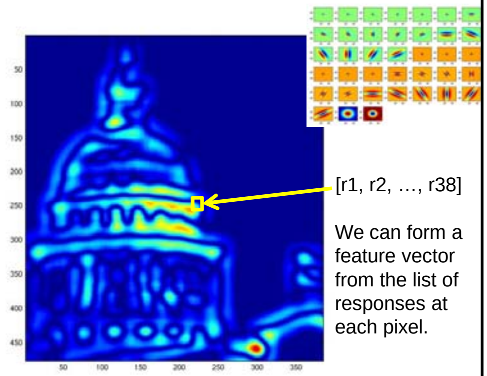
Kristen Grauman

Filter banks



- What filters to put in the bank?
  - Typically we want a combination of scales and orientations, different types of patterns.

Matlab code available for these examples:
http://www.robots.ox.ac.uk/~vgg/research/texclass/filters.html



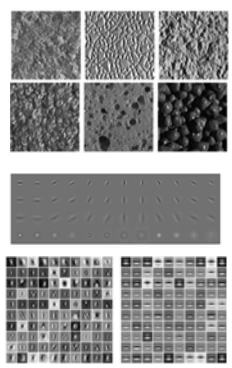
[r1, r2, …, r38]

We can form a feature vector from the list of responses at each pixel.

Kristen Grauman

Textons

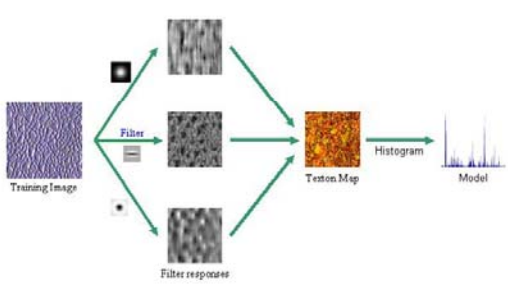
- *Texton* = cluster center of filter responses over collection of images
- Describe textures and materials based on distribution of prototypical texture elements.



Leung & Malik 1999; Varma & Zisserman, 2002

Materials as textures: example

Allows us to summarize an image according to its distribution of textons (prototypical texture patterns).



Varma & Zisserman, 2002

Manik Varma
http://www.robots.ox.ac.uk/~vgg/research/texclass/with.html

Gestalt

- Gestalt: whole or group
  - Whole is greater than sum of its parts
  - Relationships among parts can yield new properties/features
- Psychologists identified series of factors that predispose set of elements to be grouped (by human visual system)

The goals of segmentation

Separate image into coherent “objects”

image — human segmentation

Source: Lana Lazebnik

The goals of segmentation

Separate image into coherent “objects”

Group together similar-looking pixels for efficiency of further processing

“superpixels”

X. Ren and J. Malik. **Learning a classification model for segmentation.** ICCV 2003.

Source: Lana Lazebnik

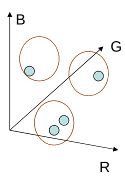
Segmentation as clustering

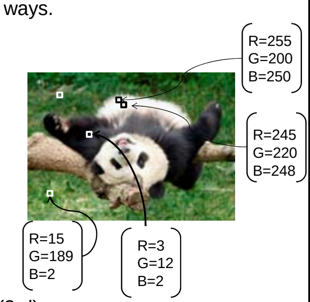
- Families of clustering algorithms
  - K-means
  - Mean shift
  - Graph cuts: normalized cuts, min-cut,…
  - Hierarchical agglomerative

Segmentation as clustering pixels

Depending on what we choose as the *feature space*, we can group pixels in different ways.

Grouping pixels based on **color** similarity

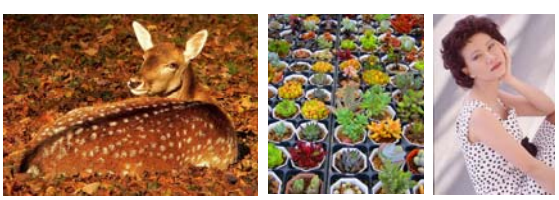




Feature space: color value (3-d)

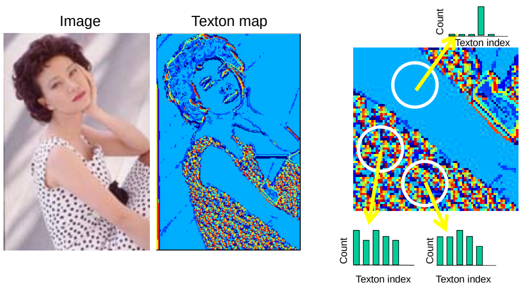
Segmentation as clustering pixels

- Color, brightness, position alone are not enough to distinguish all regions…



Segmentation with texture features

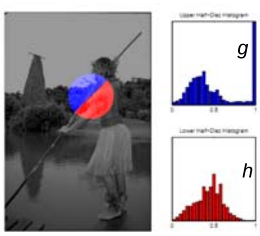
- Find "textons" by **clustering** vectors of filter bank outputs
- Describe texture in a window based on *texton histogram*



Malik, Belongie, Leung and Shi. IJCV 2001.

Adapted from Lana Lazebnik

Representing a texture gradient



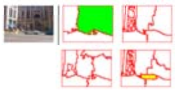
$$\chi^2(g,h) = \frac{1}{2}\sum_i \frac{(g(i)-h(i))^2}{g(i)+h(i)}$$

Figure from Arbelaez et al PAMI 2011

Ongoing topics in mid-level region representations

Multiple segmentations

- Acknowledging difficulty of finding object boundaries in **single** multi-way segmentation, now often employ **multiple segmentations** as "hypotheses"
- Input to higher-level processes.

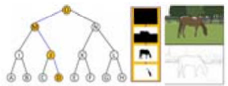


Varying parameters, grouping algorithms

Fig from Russell et al. 2006

Greedy combinations

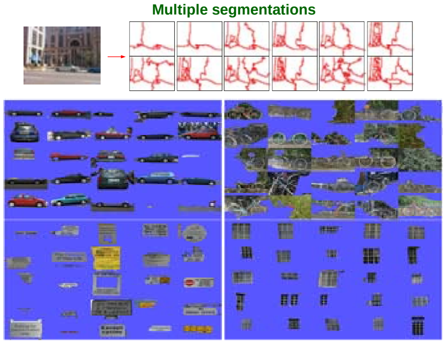
Fig from Hoiem et al. 2005



Hierarchy of segments

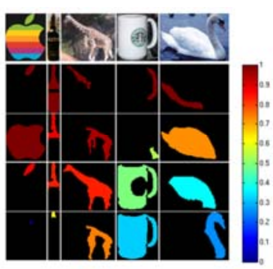
Fig from Maire et al. 2009

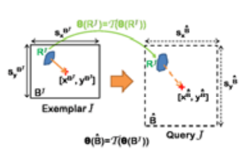
Segments as primitives for discovery



B. Russell et al., "Using Multiple Segmentations to Discover Objects and their Extent in Image Collections," CVPR 2006

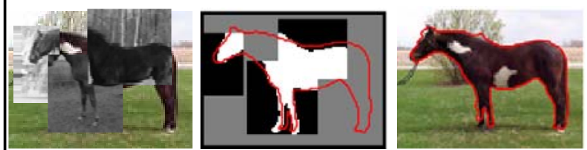
Segments as object parts





Gu et al. Recognition Using Regions, CVPR 2009

Top-down segmentation

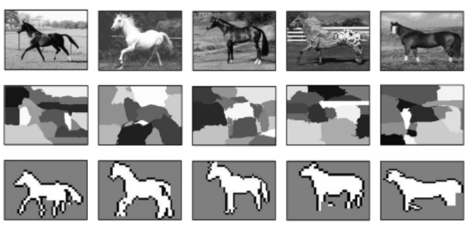


E. Borenstein and S. Ullman, "Class-specific, top-down segmentation," ECCV 2002

A. Levin and Y. Weiss, "Learning to Combine Bottom-Up and Top-Down Segmentation," ECCV 2006.

Slide credit: Lana Lazebnik

Top-down segmentation



Normalized cuts

Top-down segmentation

E. Borenstein and S. Ullman, "Class-specific, top-down segmentation," ECCV 2002

A. Levin and Y. Weiss, "Learning to Combine Bottom-Up and Top-Down Segmentation," ECCV 2006.

Slide credit: Lana Lazebnik

Motion segmentation

Input sequence | Image Segmentation | Motion Segmentation

Input sequence | Image Segmentation | Motion Segmentation

A.Barbu, S.C. Zhu. Generalizing Swendsen-Wang to sampling arbitrary posterior probabilities, *IEEE Trans. PAMI,* August 2005.

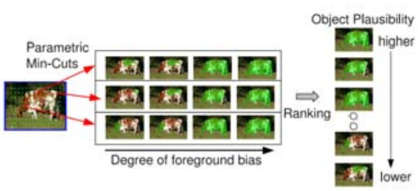
Regions to surfaces

Learn to categorize regions into geometric classes
Combining multiple segmentations

Geometric Context from a Single Image. Derek Hoiem, Alexei Efros, Martial Hebert. ICCV 2005

Category-independent ranking

How "object-like" is each candidate region?



Constrained Parametric Min-Cuts for Automatic Object Segmentation. Carreira and Sminchisescu. CVPR 2010

Also see Ferrari et al CVPR 2010, Endres et al ECCV 2010