Segmentation and localization

Kristen Grauman
UT-Austin

Announcements

- Reminder: Assignment 1 due Friday
- Assignment 2 out today, due Sept 30 and followup Oct 3

- Presenters: please send slides after class (naming instructions on website)

Today: 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

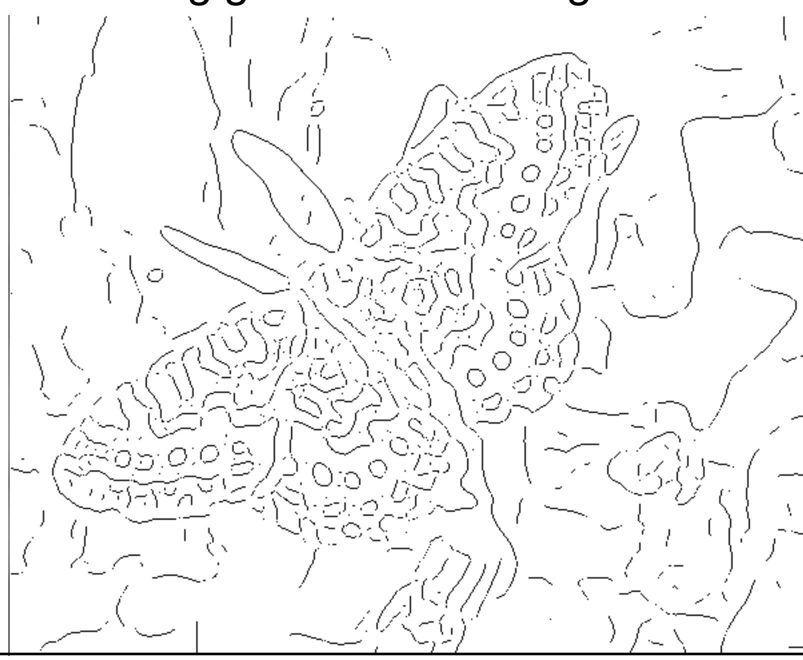
Kristen Grauman

Original image

Gradient magnitude image

Thresholding gradient with a lower threshold

Thresholding gradient with a higher threshold



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

original image (Lena)

Slide credit: Steve Seitz

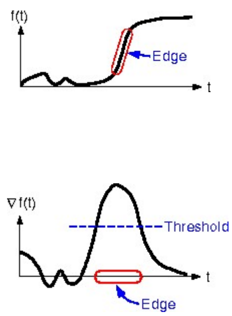
The Canny edge detector

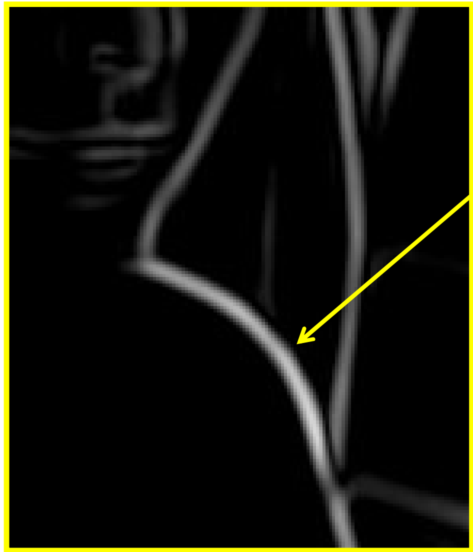
norm of the gradient

The Canny edge detector

thresholding

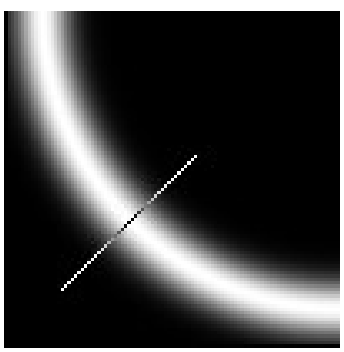
The Canny edge detector

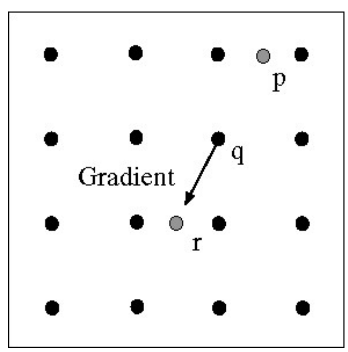




How to turn these thick regions of the gradient into curves?

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

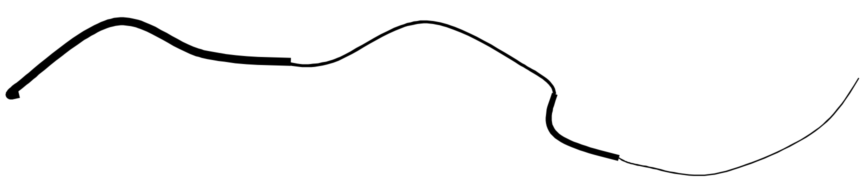
The Canny edge detector

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

thinning
(non-maximum suppression)

Hysteresis thresholding

- Use a high threshold to start edge curves, and a low threshold to continue them.



Source: Steve Seitz

Hysteresis thresholding

original image

high threshold (strong edges)

low threshold (weak edges)

hysteresis threshold

Source: L. Fei-Fei

Hysteresis thresholding

high threshold (strong edges)

low threshold (weak edges)

hysteresis threshold

Source: L. Fei-Fei

Recap: 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

Low-level edges vs. perceived contours

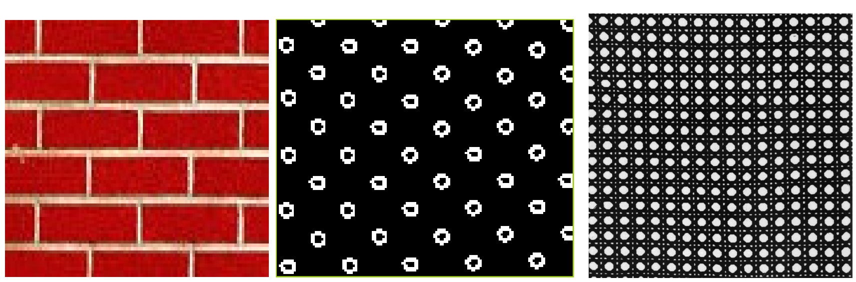
Background **Shadows** **Texture**

Kristen Grauman

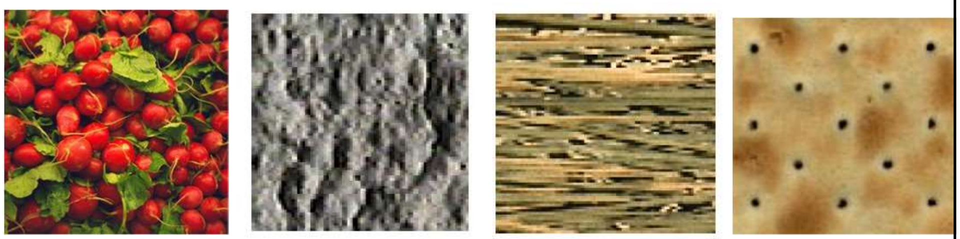
Texture

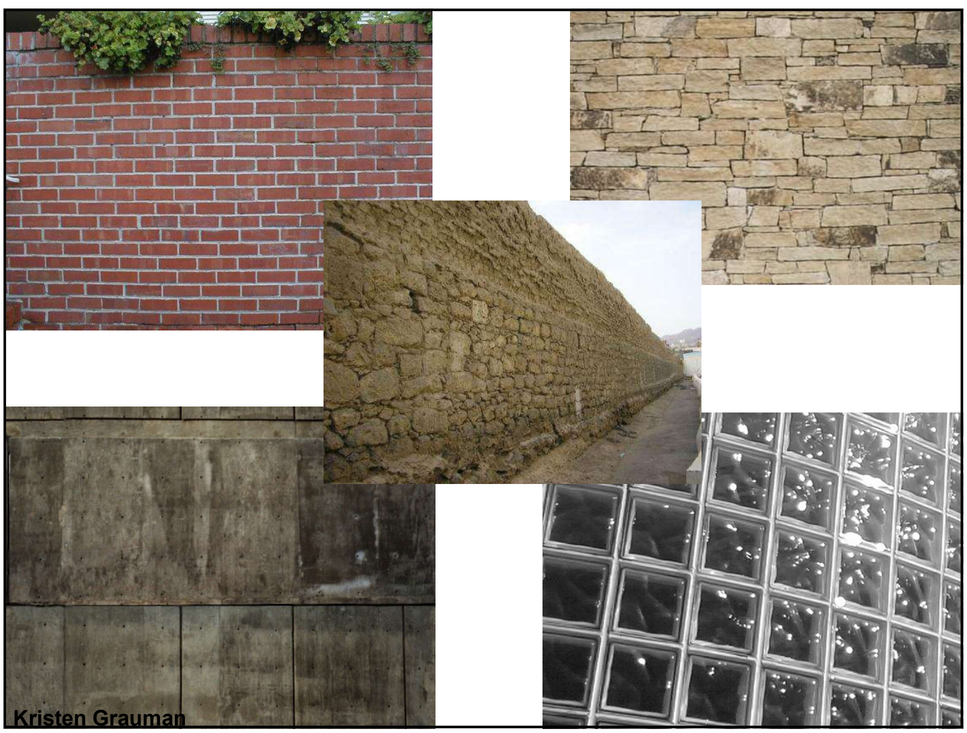
What defines a texture?

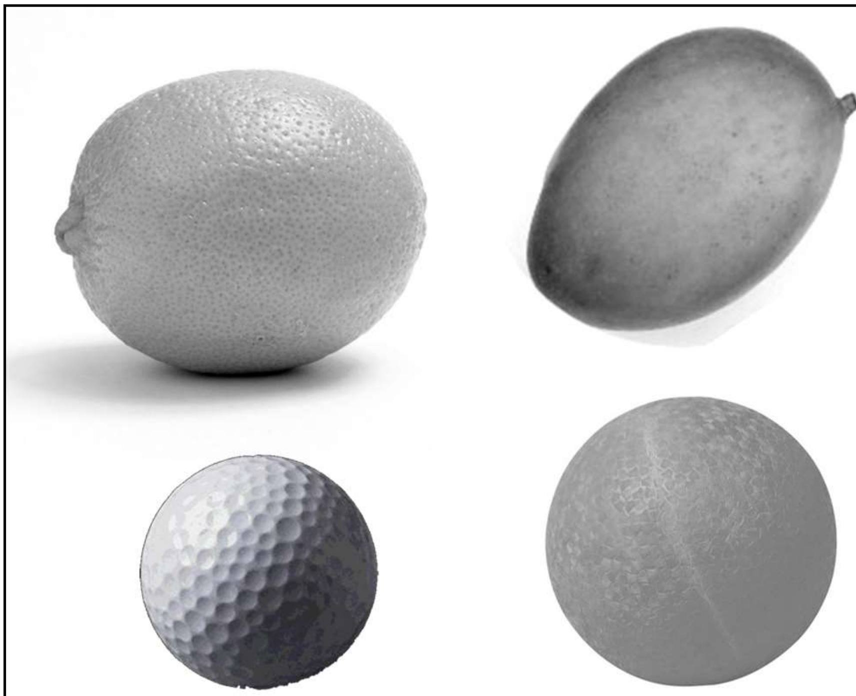
Includes: more regular patterns



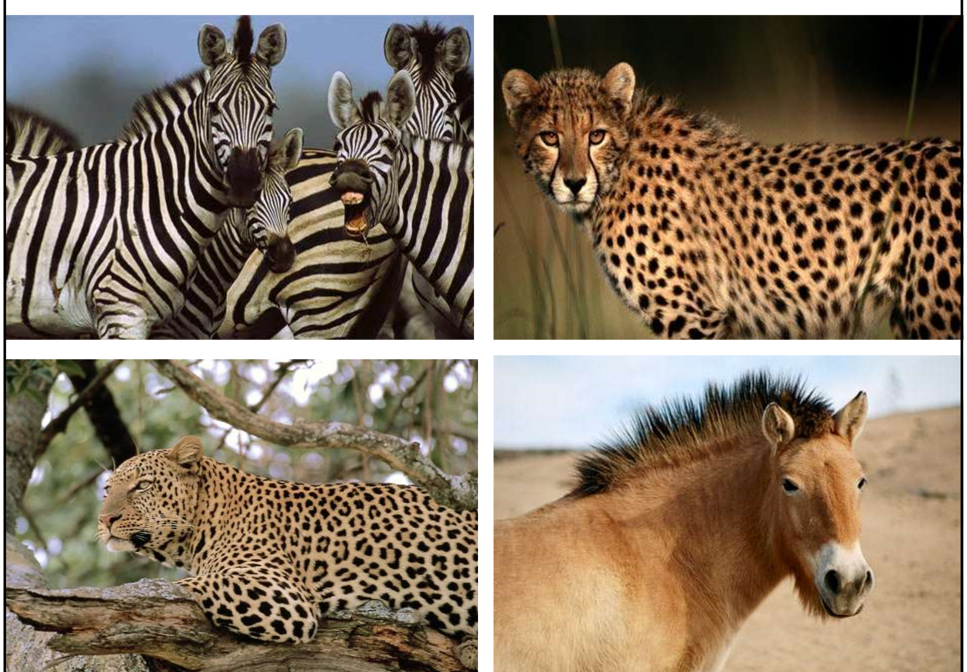
Includes: more random patterns







Kristen Grauman



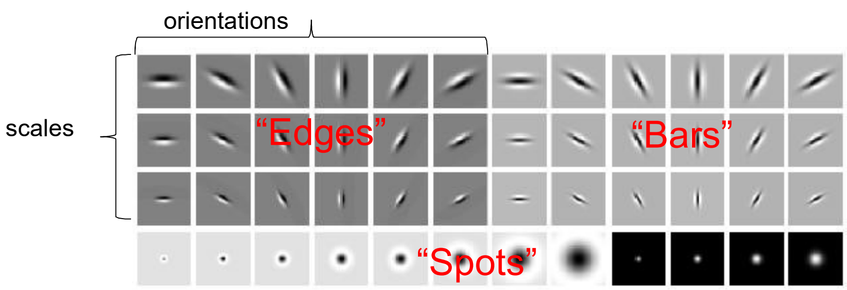
Kristen Grauman http://animals.nationalgeographic.com/

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

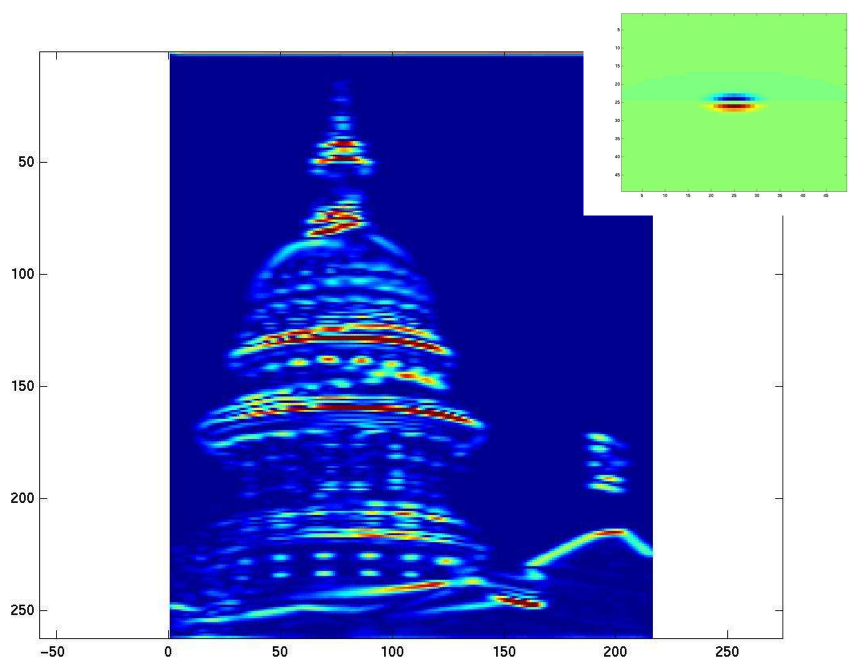


- What filters to put in the bank?
  - Typically we want a combination of scales and orientations, different types of patterns.

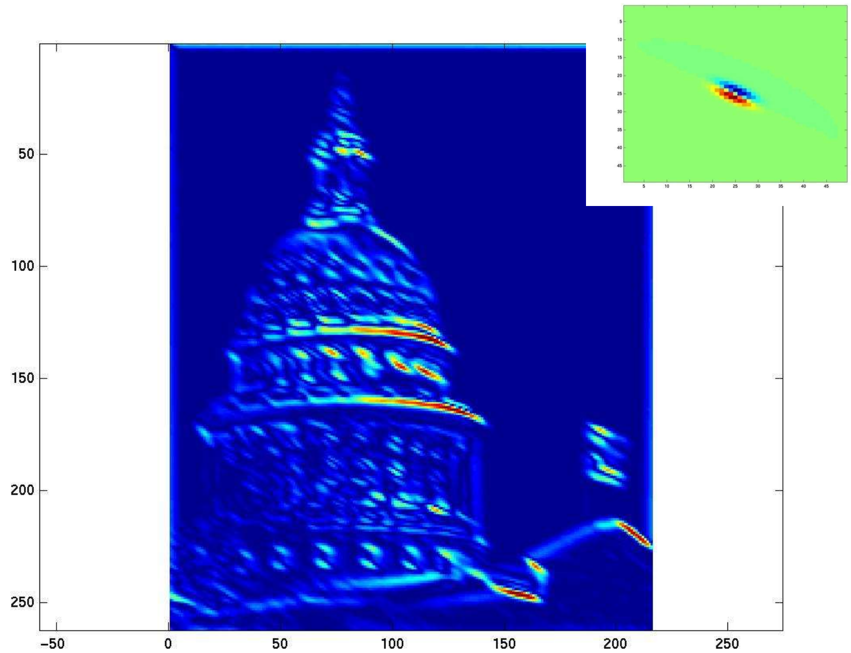
Matlab code available for these examples:
http://www.robots.ox.ac.uk/~vgg/research/texclass/filters.html

Image from http://www.texasexplorer.com/austincap2.jpg

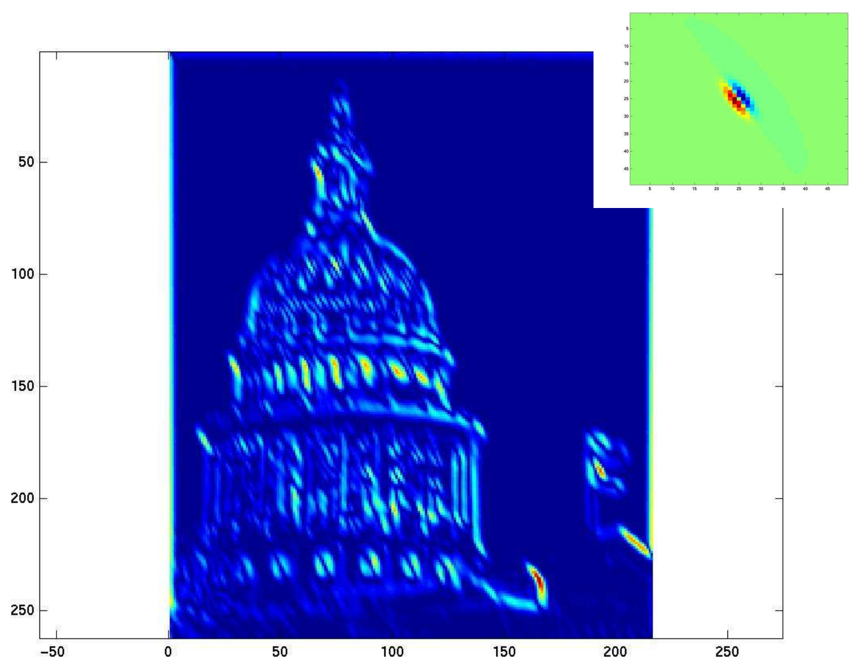
Kristen Grauman



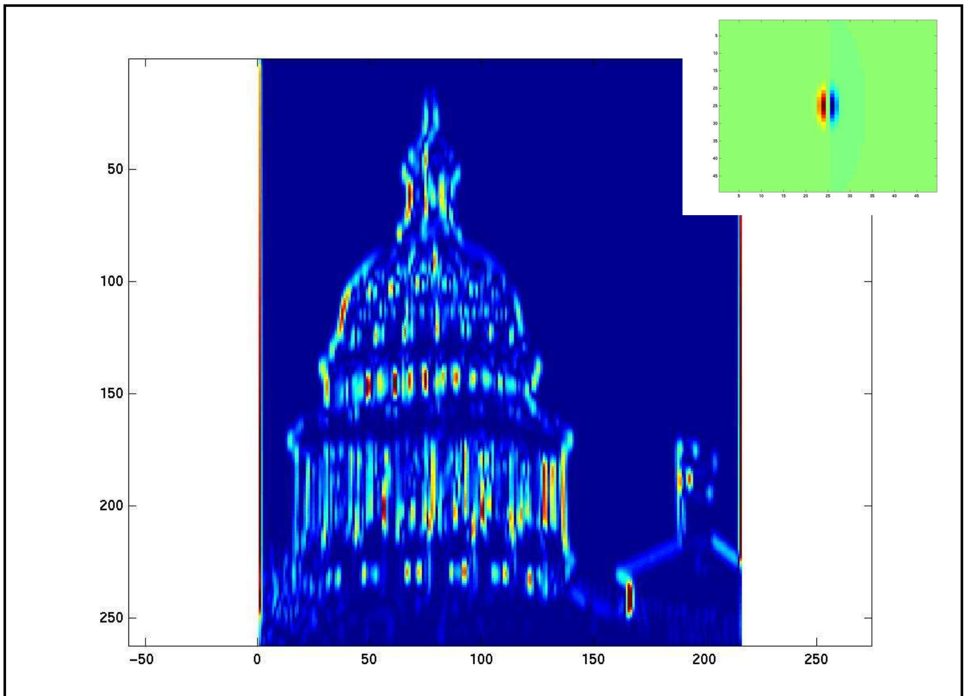
Kristen Grauman



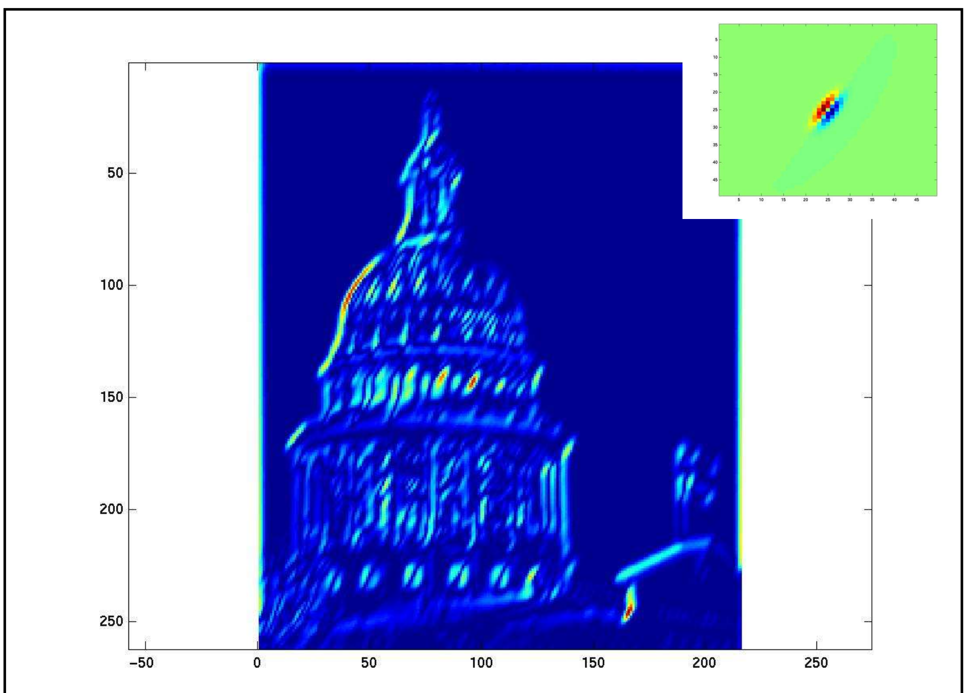
Kristen Grauman



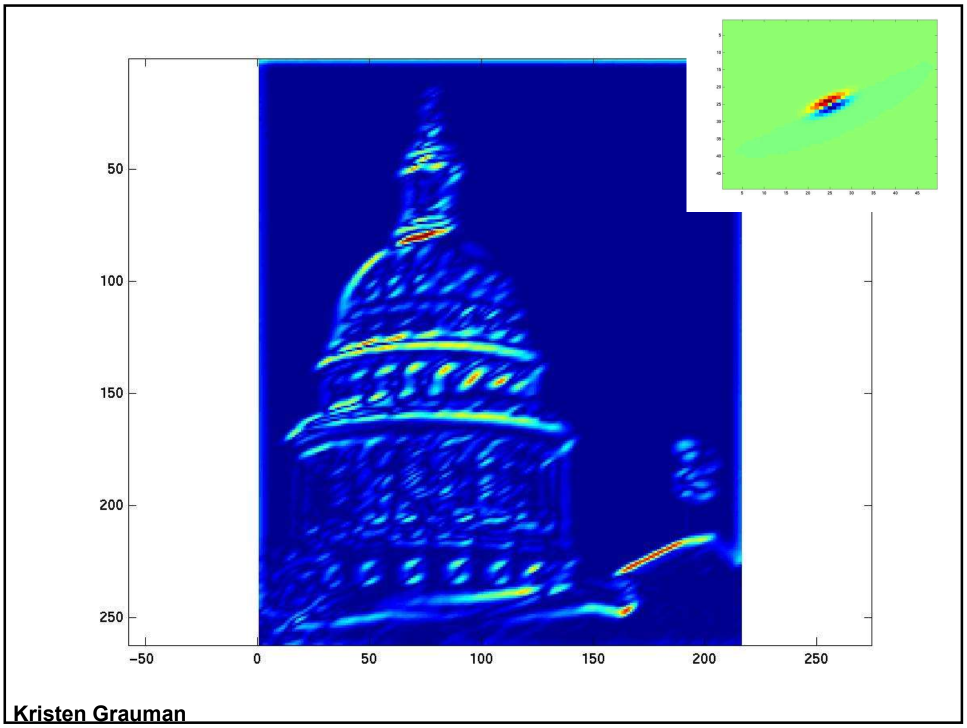
Kristen Grauman

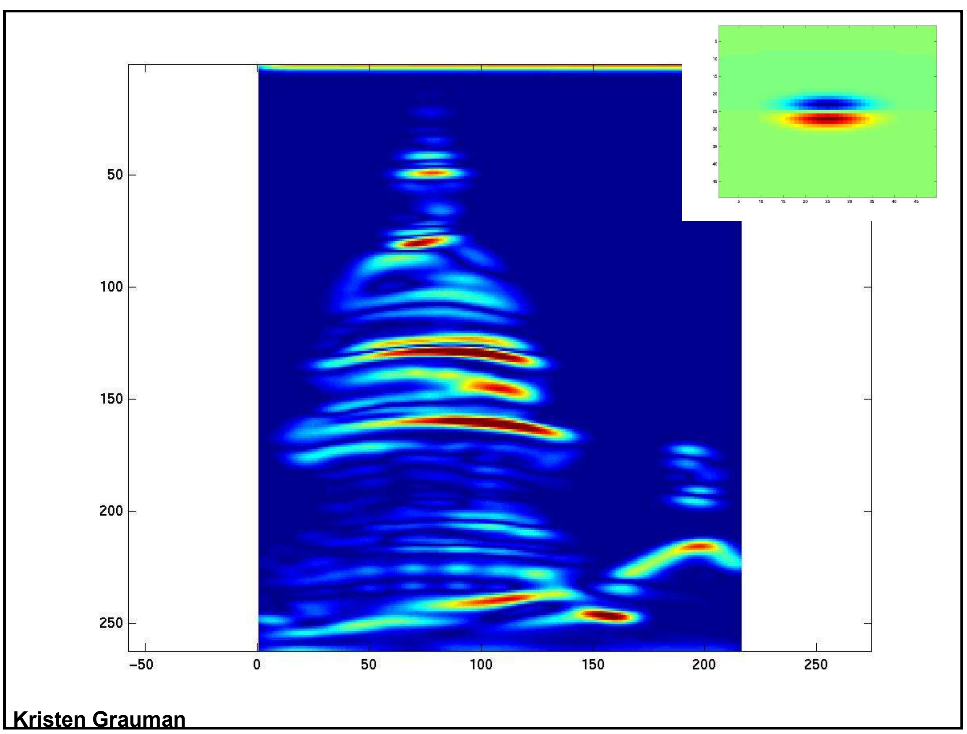


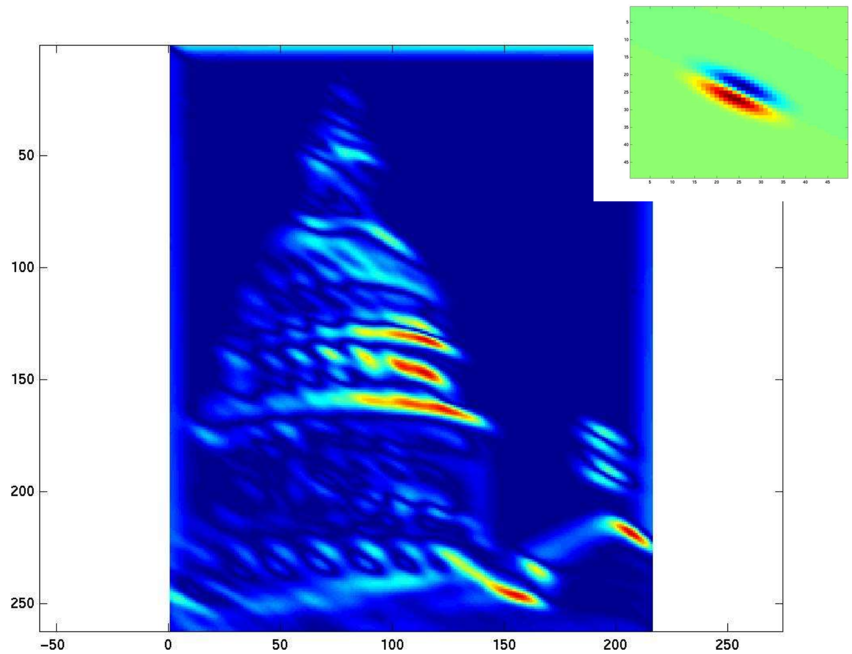
Kristen Grauman



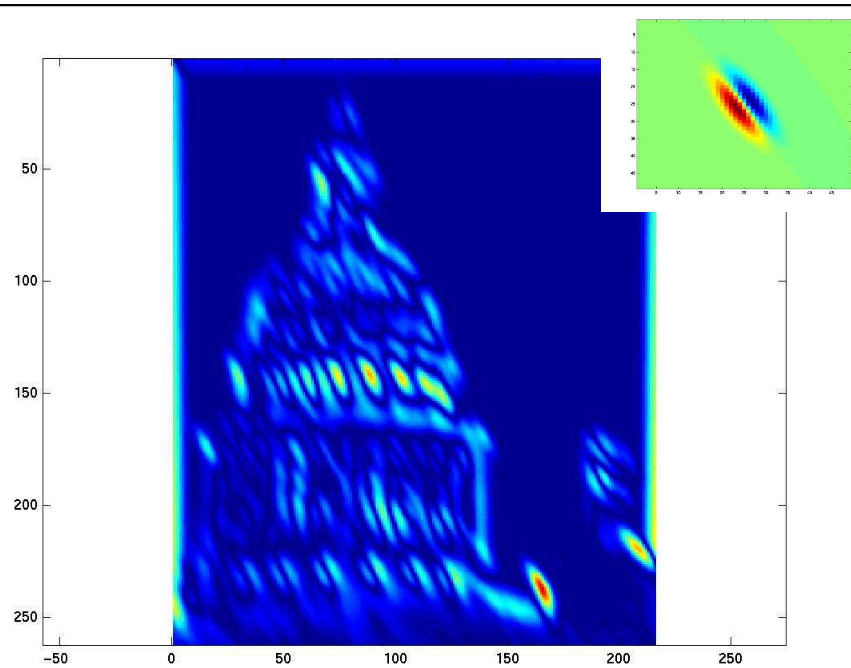
Kristen Grauman


50
100
150
200
250
−50
0
50
100
150
200
250
Kristen Grauman

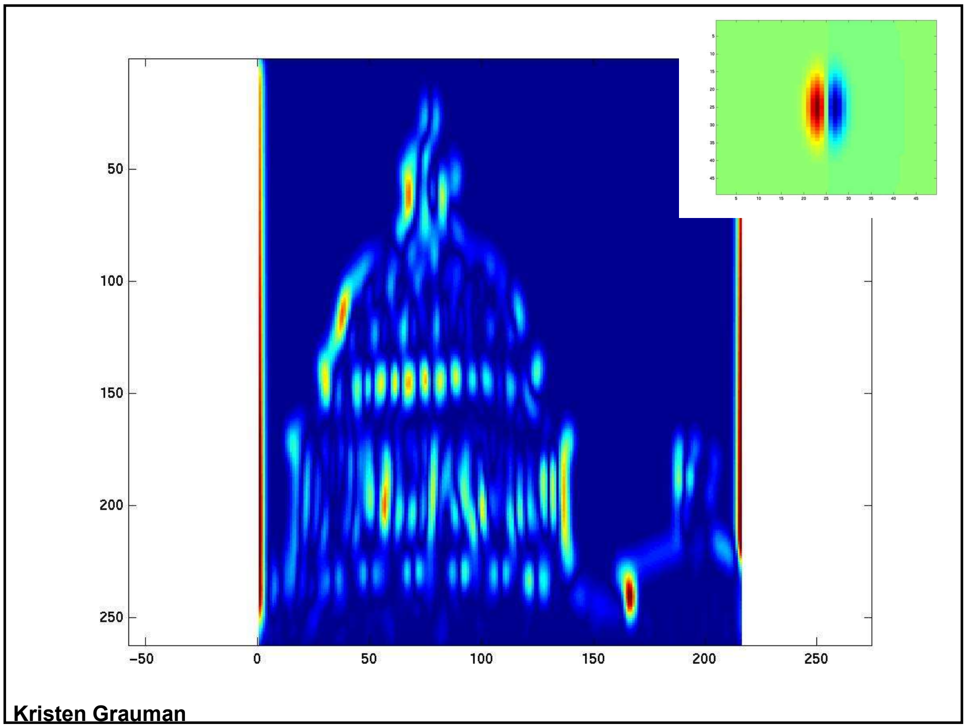

50
100
150
200
250
−50
0
50
100
150
200
250
Kristen Grauman

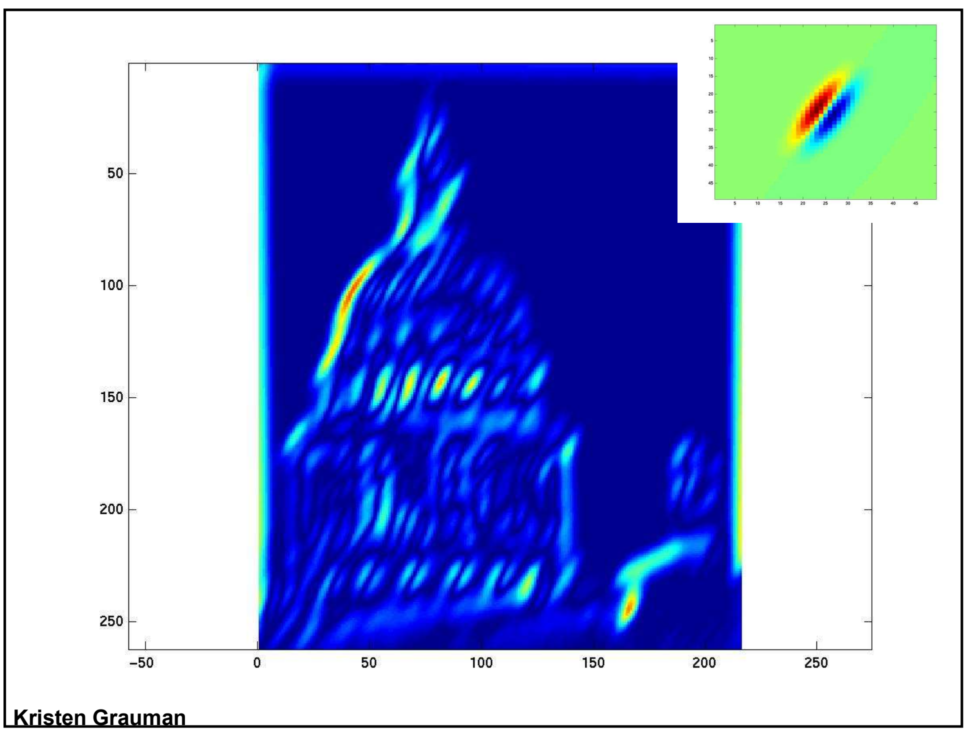


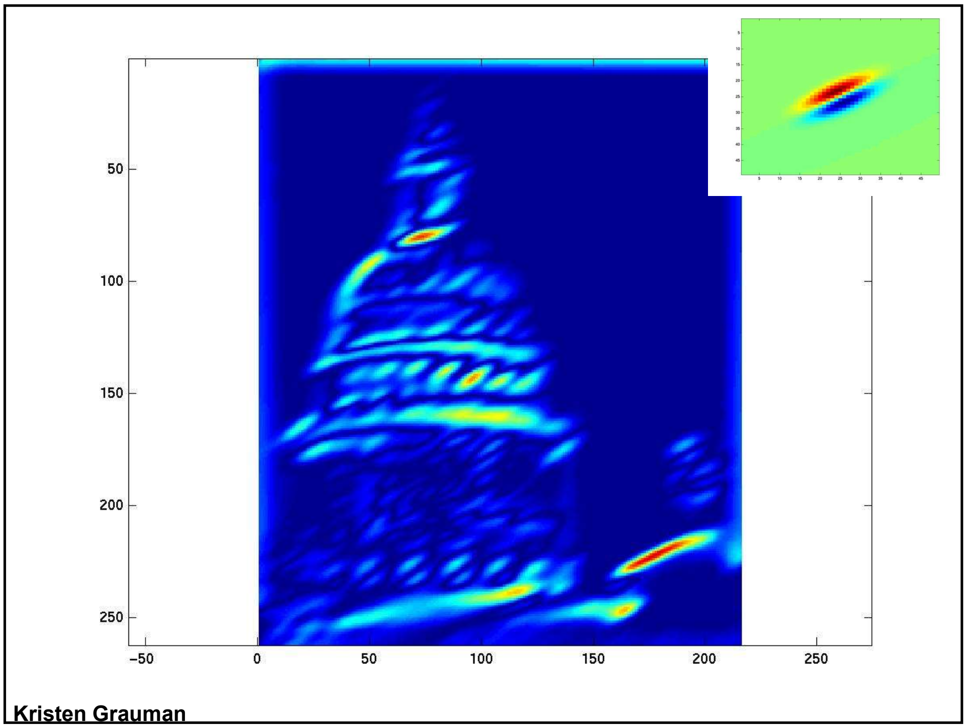
Kristen Grauman



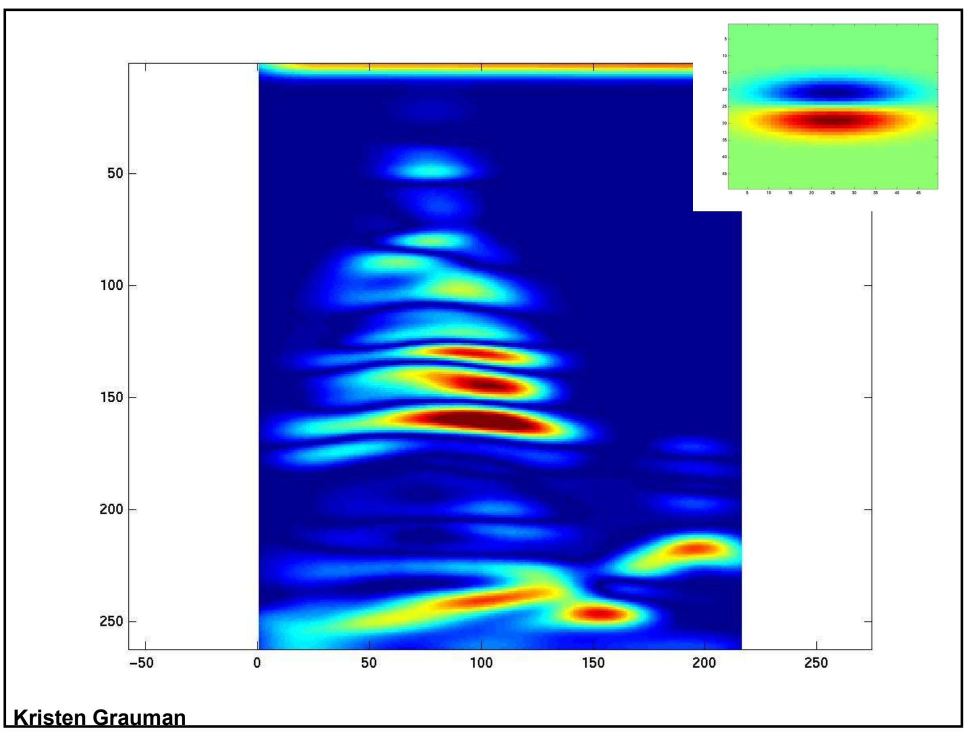
Kristen Grauman


50
100
150
200
250
−50
0
50
100
150
200
250
Kristen Grauman

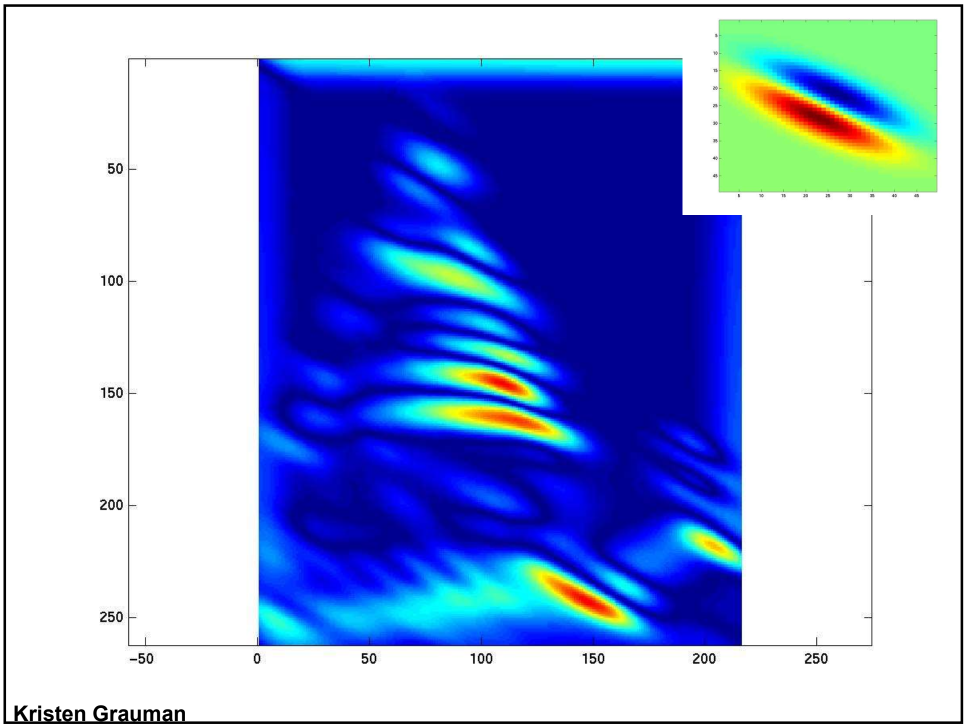

50
100
150
200
250
−50
0
50
100
150
200
250
Kristen Grauman

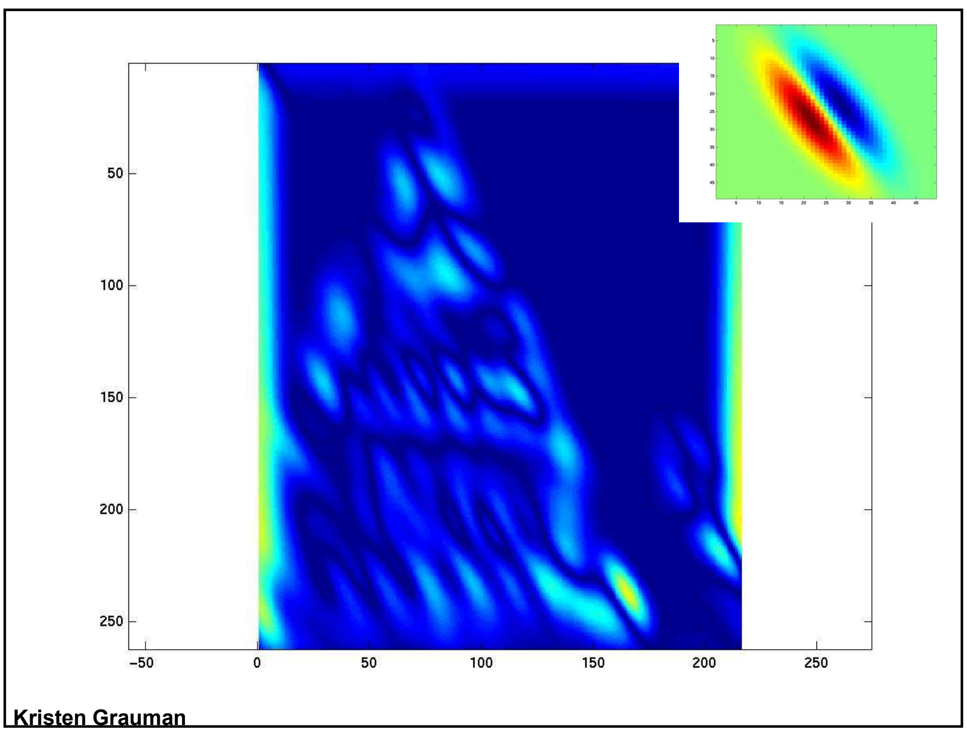


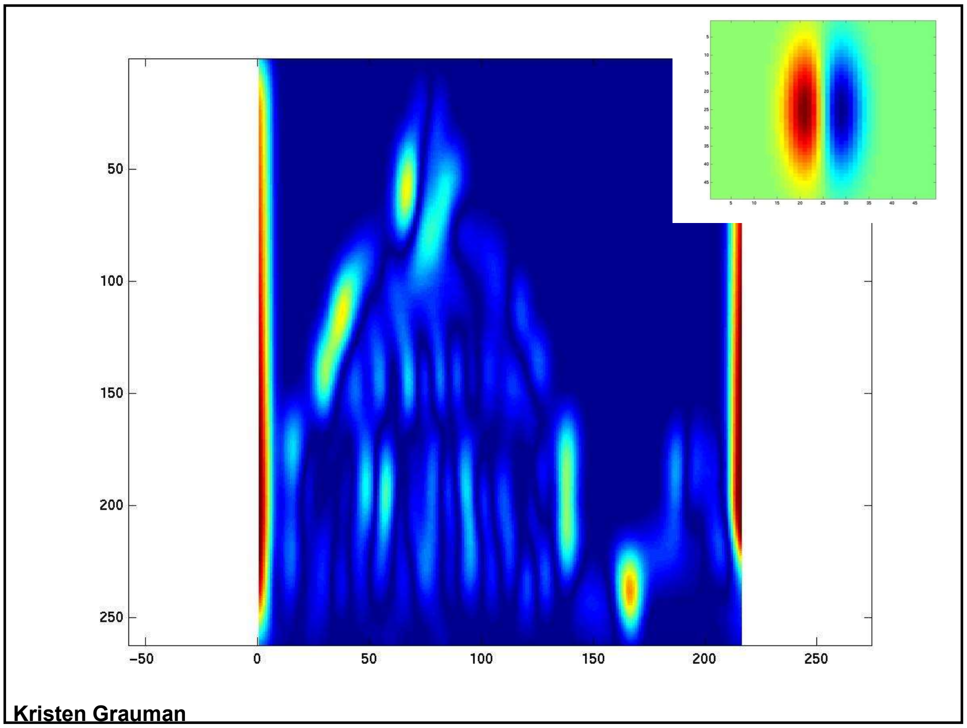
50
100
150
200
250
−50
0
50
100
150
200
250
Kristen Grauman



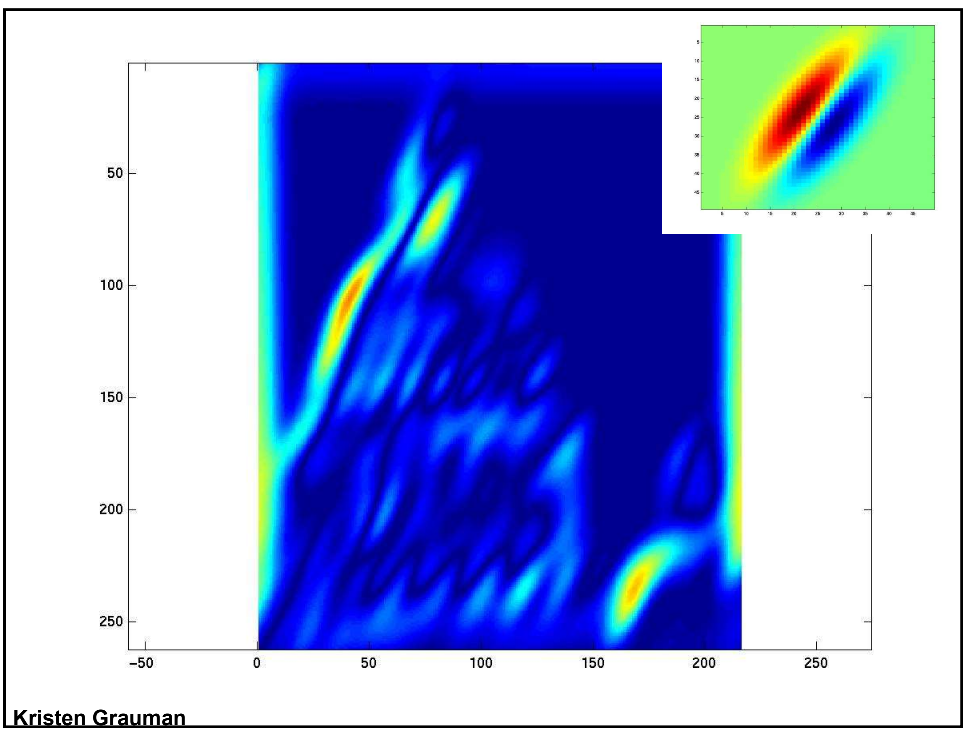
50
100
150
200
250
−50
0
50
100
150
200
250
Kristen Grauman


50
100
150
200
250
−50
0
50
100
150
200
250
Kristen Grauman

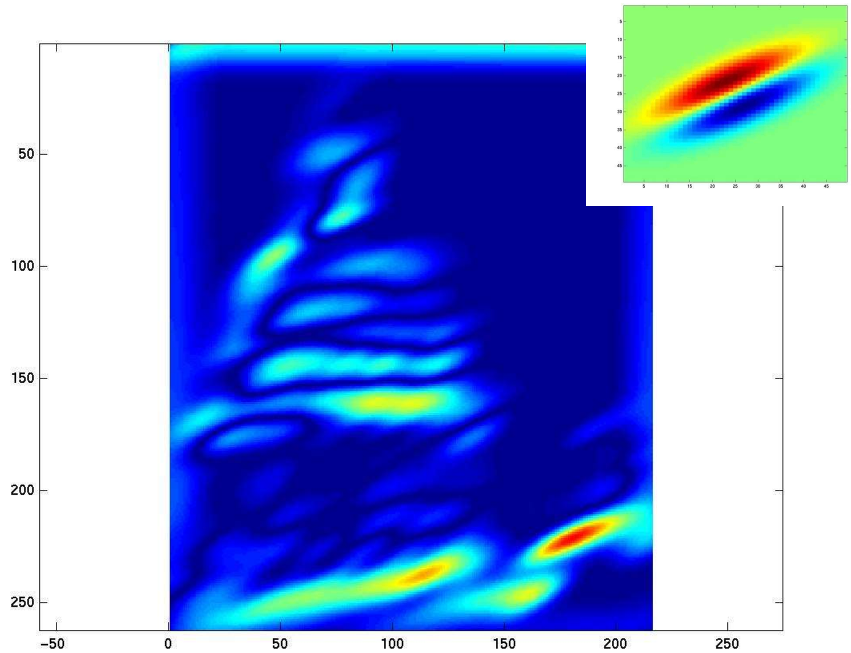

50
100
150
200
250
−50
0
50
100
150
200
250
Kristen Grauman



Kristen Grauman



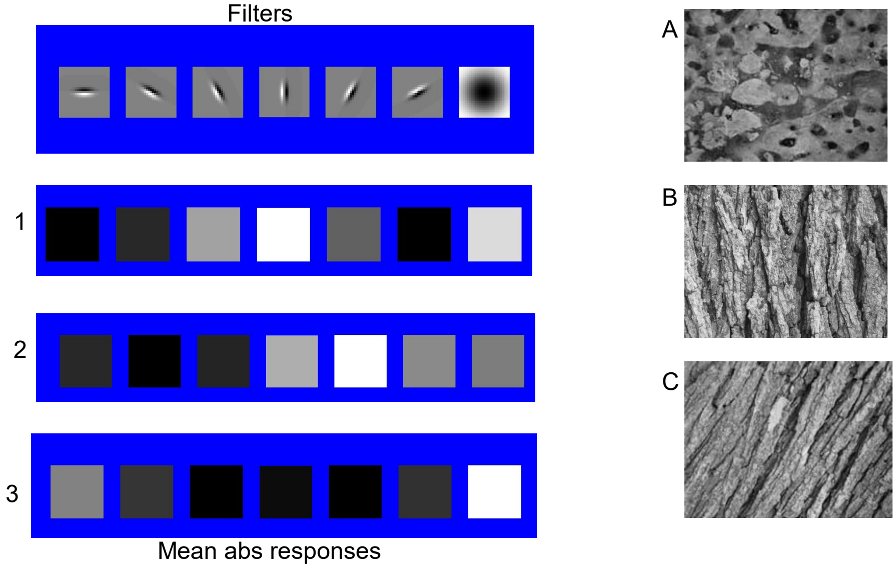
Kristen Grauman



Kristen Grauman

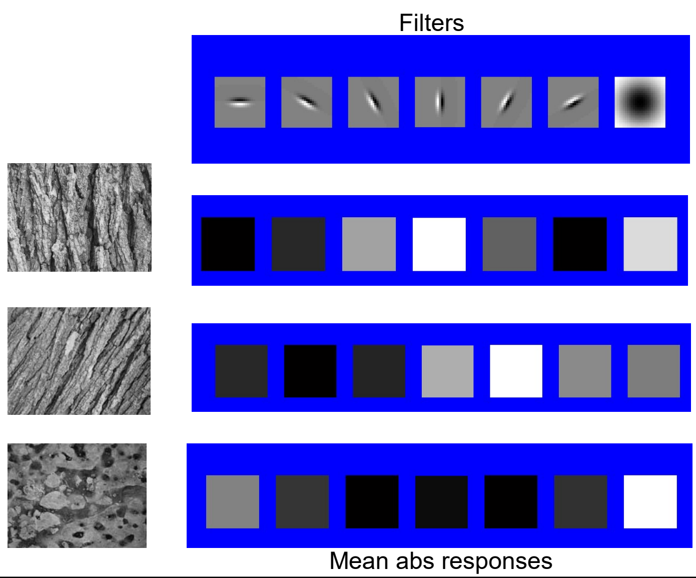
Kristen Grauman

You try: Can you match the texture to the response?

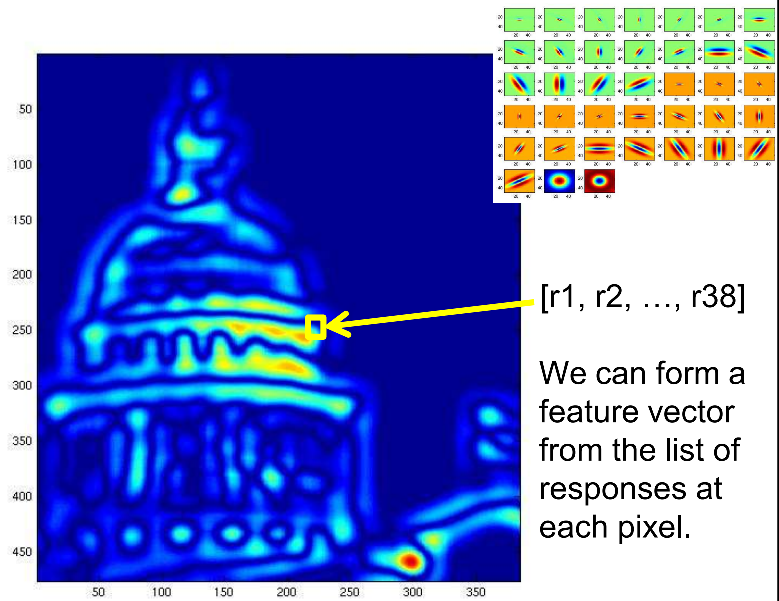


Derek Hoiem

Representing texture by mean abs response



Derek Hoiem



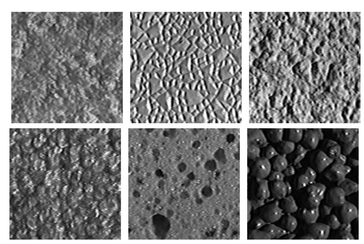
[r1, r2, …, r38]

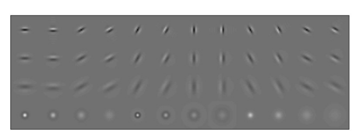
We can form a feature vector from the list of responses at each pixel.

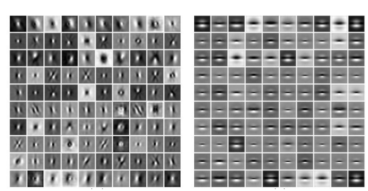
Kristen Grauman

Textons

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



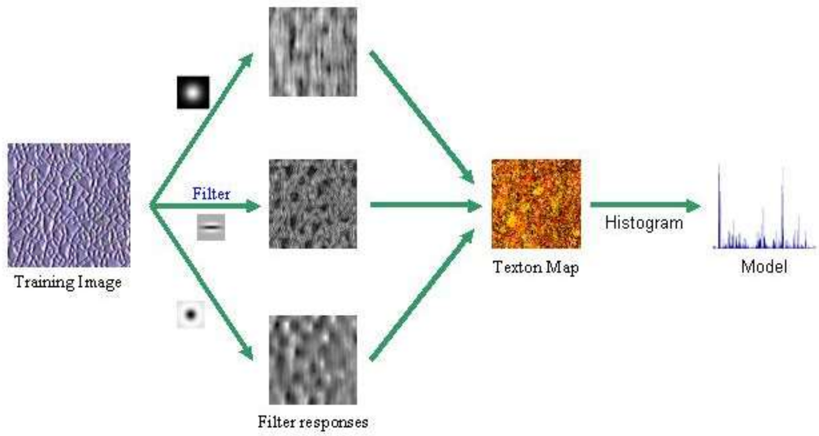




Leung & Malik 1999; Varma & Zisserman, 2002

Materials as textures: example

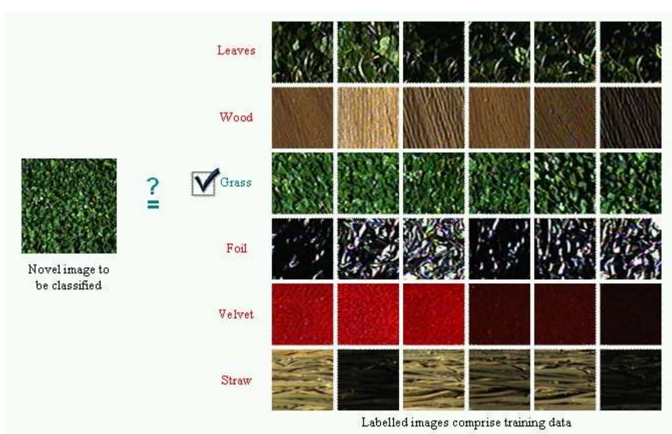
Allows us to summarize an image according to its distribution of textons (prototypical texture patterns).



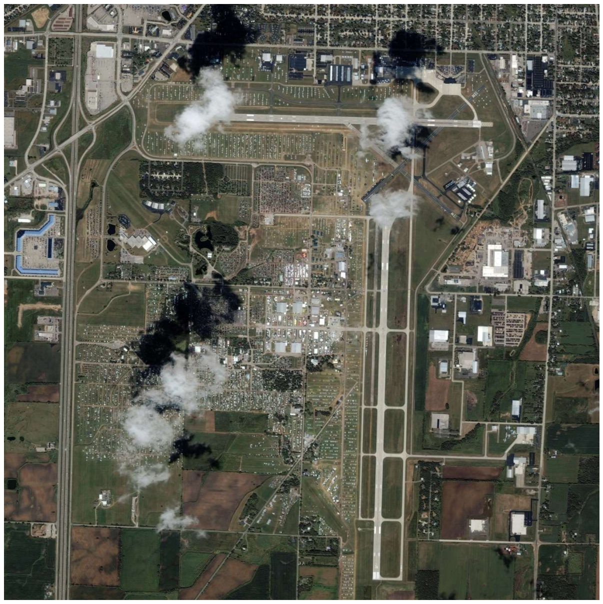
Varma & Zisserman, 2002

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

Materials as textures: example



Varma & Zisserman, 2002

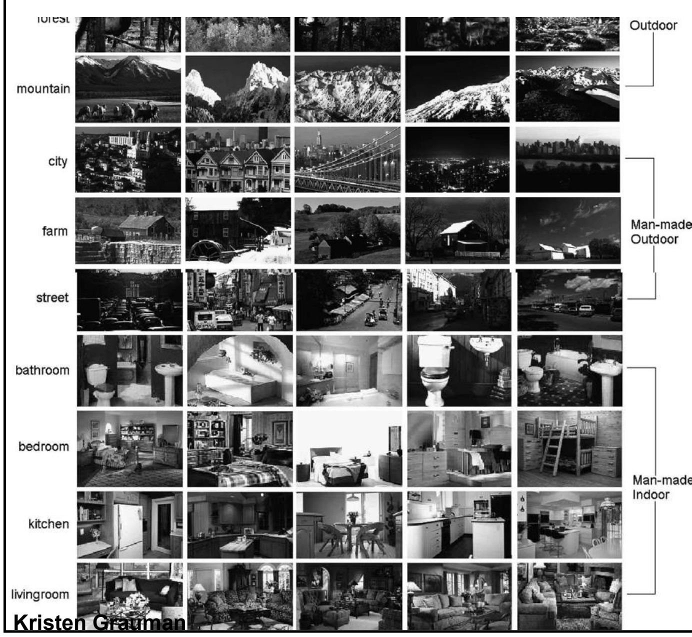


Segmenting aerial imagery by textures

Kristen Grauman

http://www.airventure.org/2004/gallery/images/073104_satellite.jpg

Scenes as textures: example



Characterizing scene categories by texture

L. W. Renninger and J. Malik. When is scene identification just texture recognition? Vision Research 44 (2004) 2301–2311

Kristen Grauman

Texture: recap

- Texture is a useful property that is often indicative of materials, appearance cues
- **Texture representations** attempt to summarize repeating patterns of local structure
- **Filter banks** useful to measure redundant variety of structures in local neighborhood

Kristen Grauman

Mid-level cues

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

- Edges, contours
- Texture
- Regions
- Surfaces

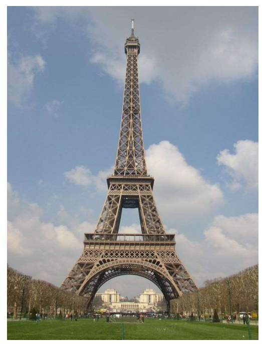
Gestalt

- 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)

Similarity

http://chicagoist.com/attachments/chicagoist_alicia/GEESE.jpg, http://wwwdelivery.superstock.com/WI/223/1532/PreviewComp/SuperStock_1532R-0831.jpg

Symmetry



http://seedmagazine.com/news/2006/10/beauty_is_in_the_processingtim.php

Common fate

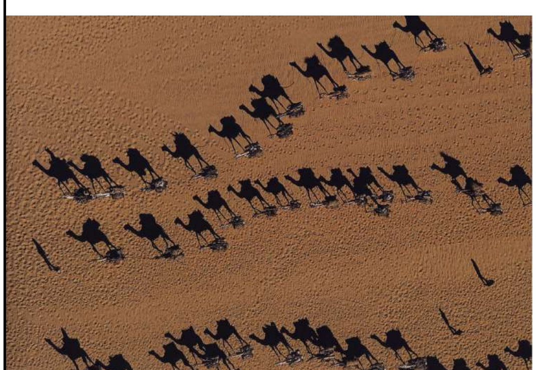
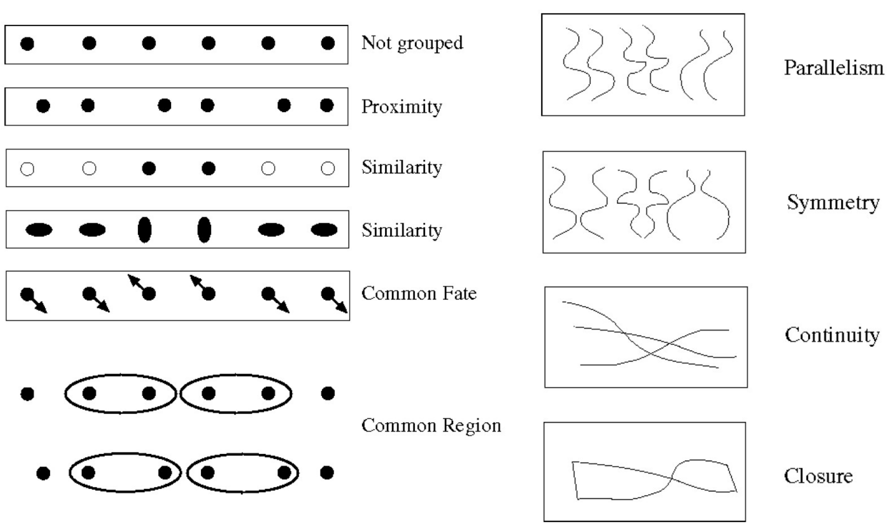


Image credit: Arthus-Bertrand (via F. Durand)

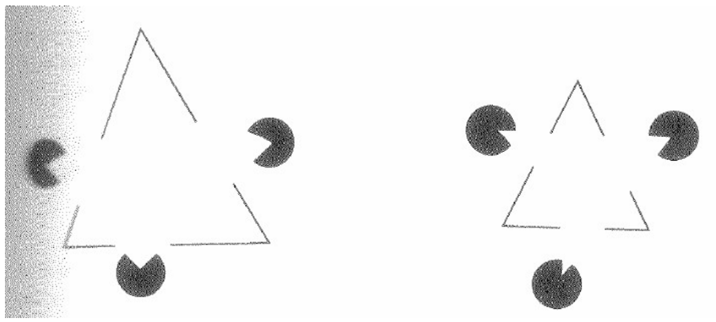
Proximity

http://www.capital.edu/Resources/Images/outside6_035.jpg

Some Gestalt factors

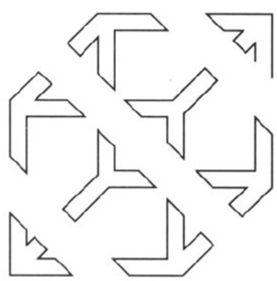


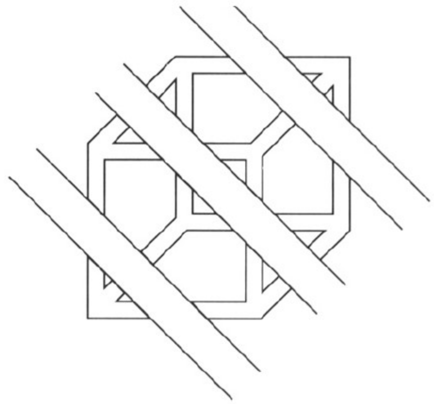
Illusory/subjective contours



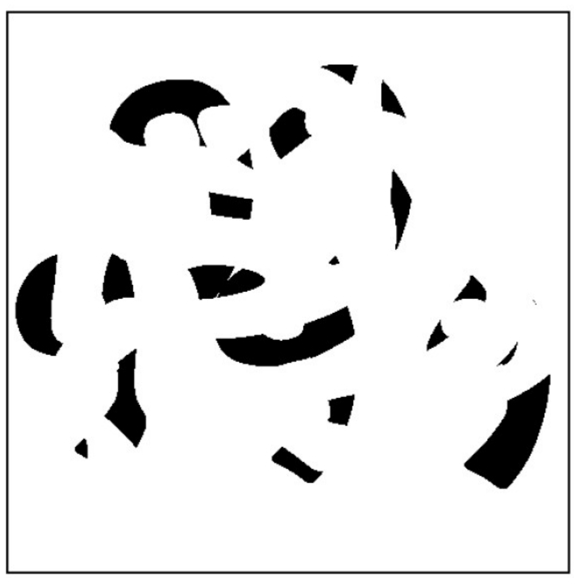
Interesting tendency to explain by occlusion

In *Vision*, D. Marr, 1982





Continuity, explanation by occlusion



D. Forsyth

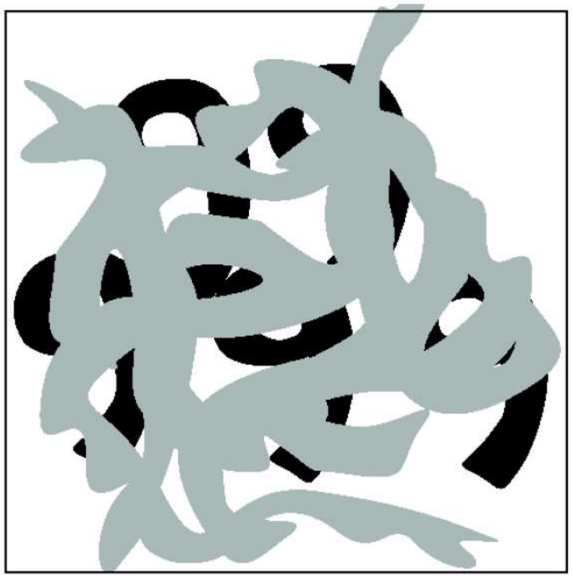
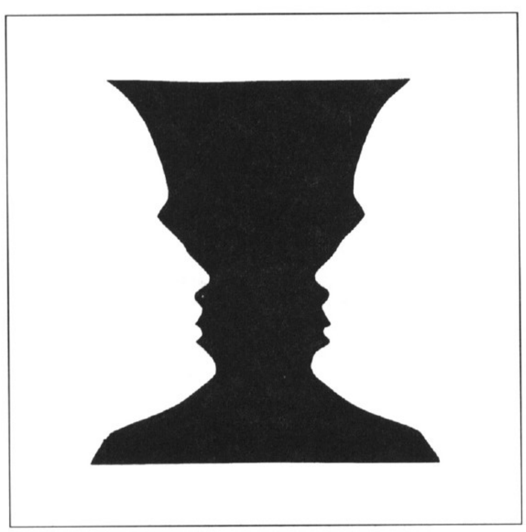
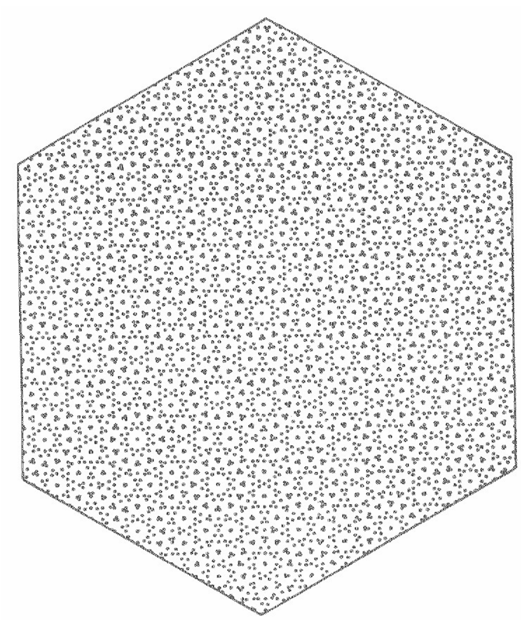


Figure-ground



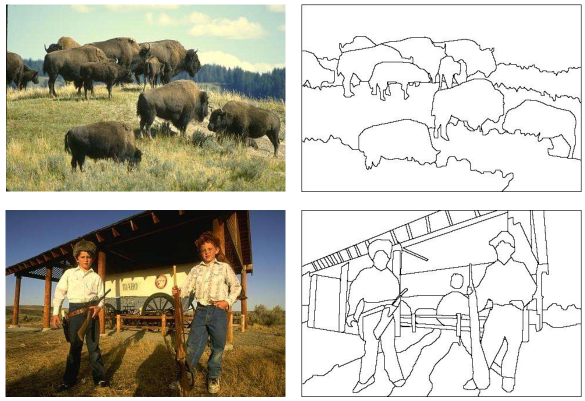


In *Vision*, D. Marr, 1982; from J. L. Marroquin, “Human visual perception of structure”, 1976.

The goals of segmentation

Separate image into coherent "objects"

image | human segmentation



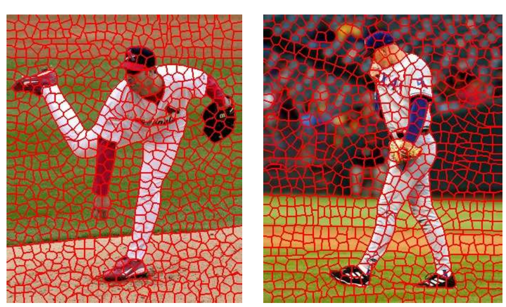
Source: Lana Lazebnik

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

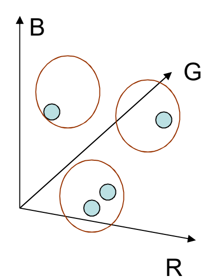
Segmentation as clustering

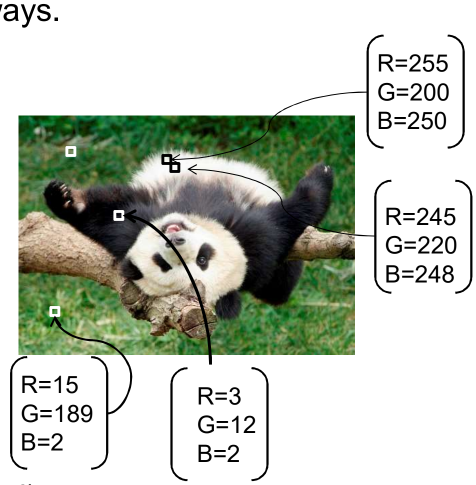
- Families of clustering algorithms
  - K-means
  - Mean shift
  - Graph cuts: normalized cuts, min-cut,…
  - Hierarchical agglomerative

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*

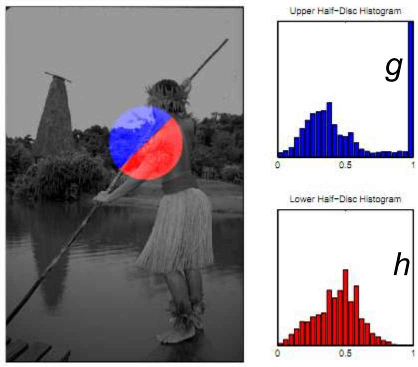
Image

Texton map

Malik, Belongie, Leung and Shi. IJCV 2001.

Adapted from Lana Lazebnik

Representing a texture gradient

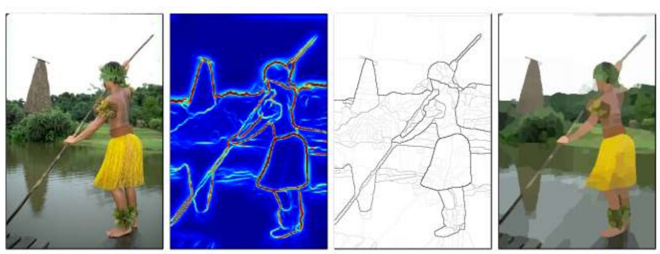


$$\chi^2(g,h) = \frac{1}{2}\sum_i \frac{(g(i)-h(i))^2}{g(i)+h(i)}$$

Figure from Arbelaez et al PAMI 2011

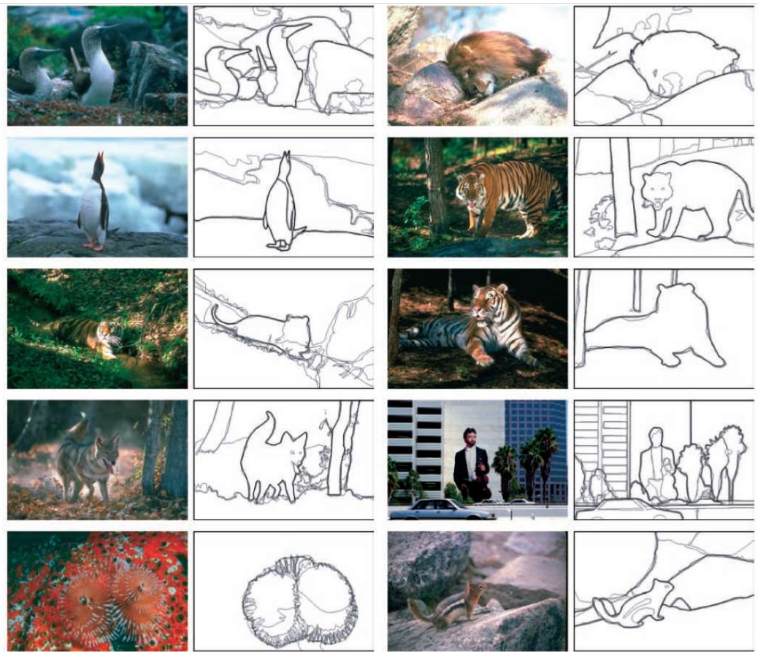
Contour Detection and Hierarchical Image Segmentation

Pablo Arbelaez, Michael Maire, Charless Fowlkes, Jitendra Malik



Predict contours based on oriented gradients
Map to closed regions with watershed
Hierarchy of segments as output

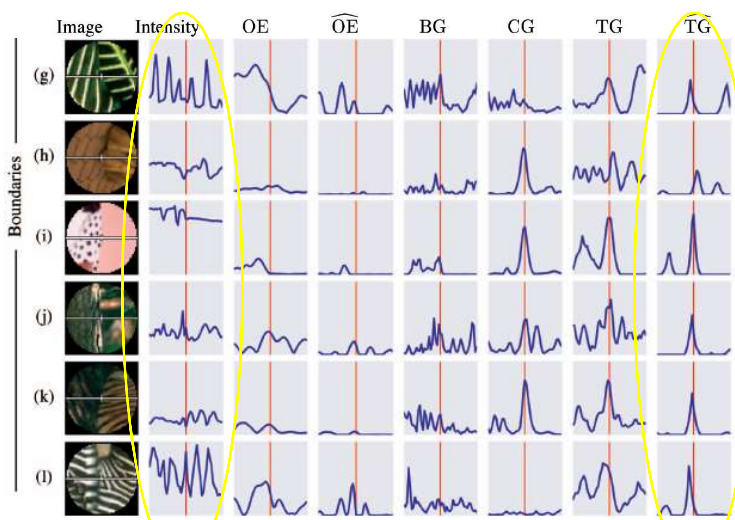
Idea: learn from humans which combination of features is most indicative of a "good" contour



[D. Martin et al. PAMI 2004]

Human-marked segment boundaries

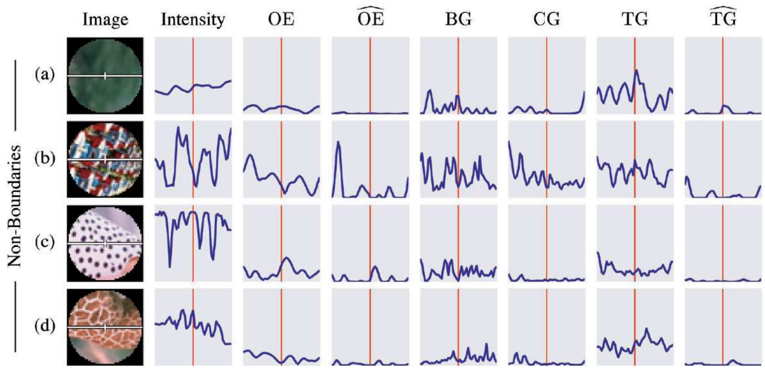
What features are responsible for perceived edges?



Feature profiles (oriented energy, brightness, color, and texture gradients) along the patch's horizontal diameter

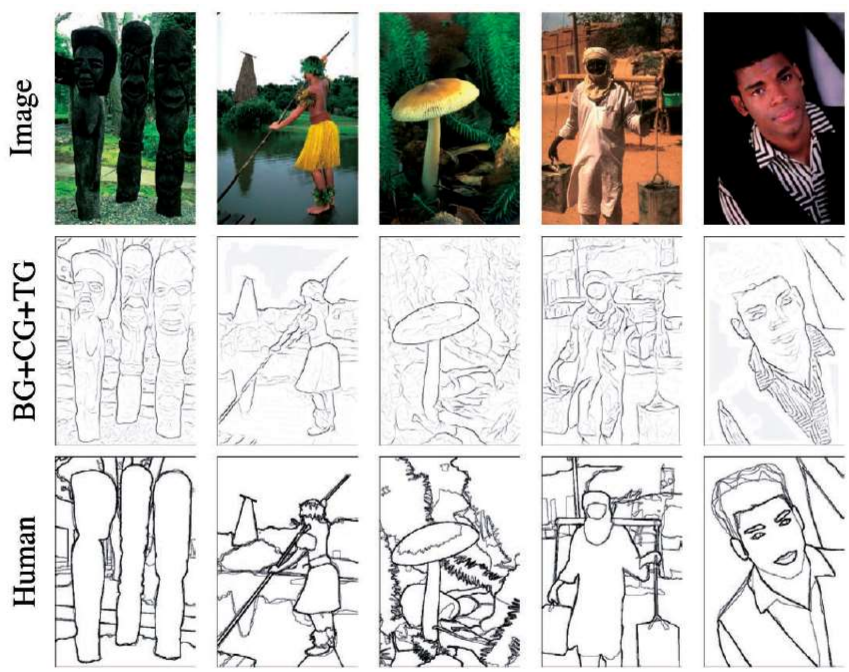
[D. Martin et al. PAMI 2004]

What features are responsible for perceived edges?



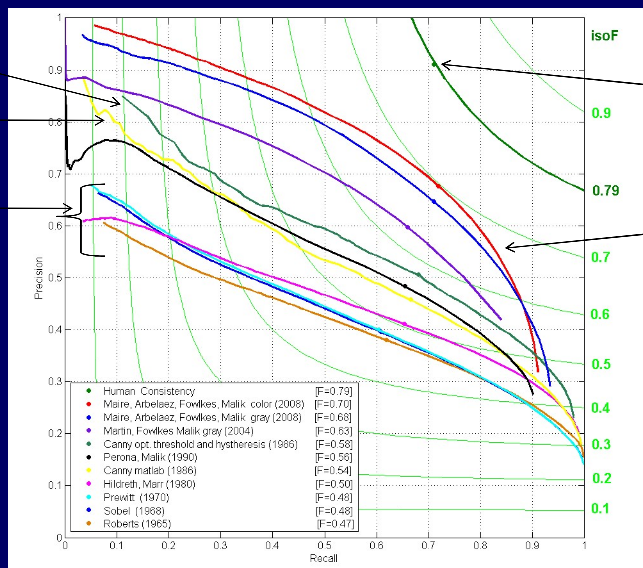
Feature profiles (oriented energy, brightness, color, and texture gradients) along the patch's horizontal diameter

[D. Martin et al. PAMI 2004]



[D. Martin et al. PAMI 2004]

Contour Detection



Canny+opt thresholds

Canny

Prewitt, Sobel, Roberts

Human agreement

Learned with combined features

UC Berkeley

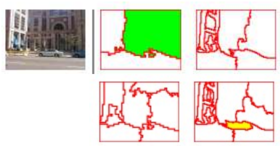
Source: Jitendra Malik: http://www.cs.berkeley.edu/~malik/malik-talks-ptrs.html

Computer Vision Group

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.

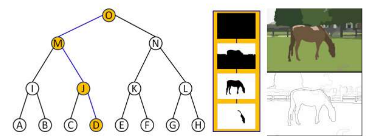


Varying parameters, grouping algorithms

Fig from Russell et al. 2006

Greedy combinations

Fig from Hoiem et al. 2005

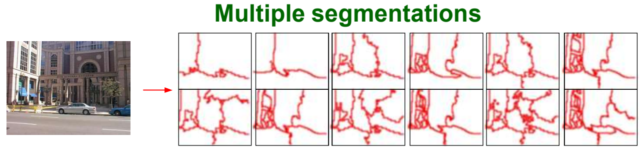


Hierarchy of segments

Fig from Maire et al. 2009

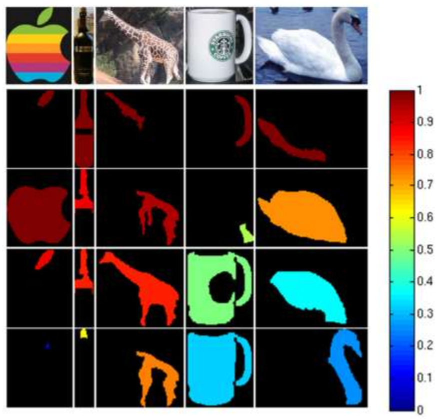
Segments as primitives for discovery

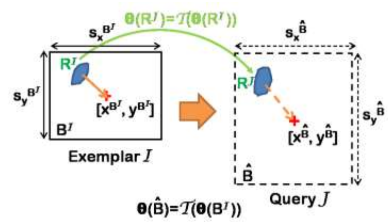
Multiple segmentations



B. Russell et al., [“Using Multiple Segmentations to Discover Objects and their Extent in Image Collections,”]() CVPR 2006

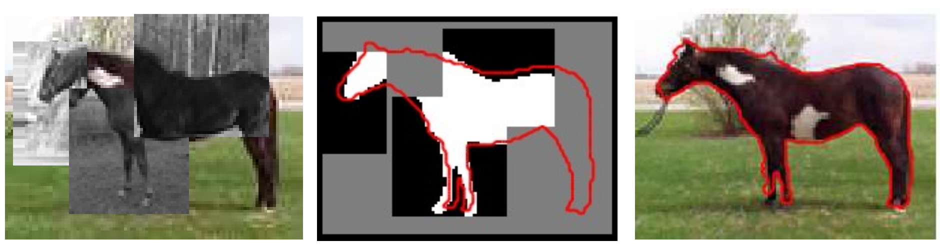
Segments as object parts?





Gu et al. Recognition Using Regions, CVPR 2009

Top-down segmentation

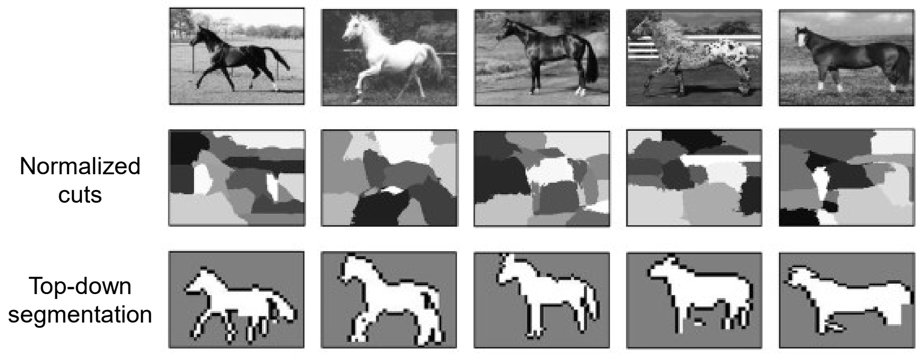


E. Borenstein and S. Ullman, "Class-specific, top-down segmentation," ECCV 2002

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

Slide credit: Lana Lazebnik

Top-down segmentation

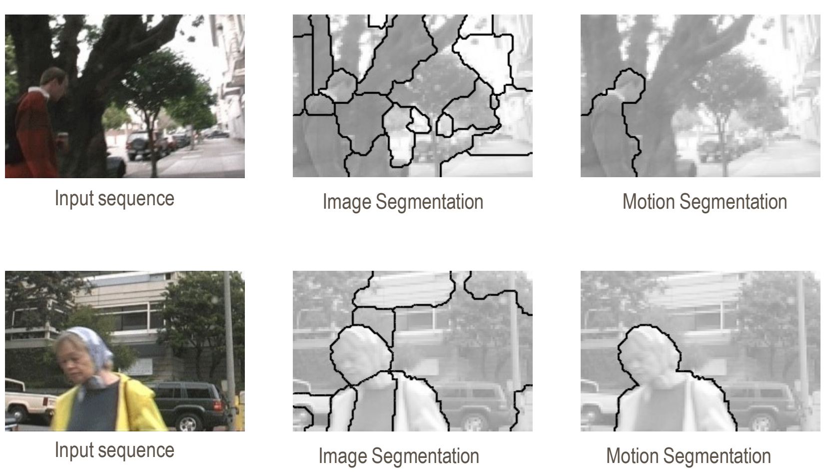


E. Borenstein and S. Ullman, "Class-specific, top-down segmentation," ECCV 2002

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

Slide credit: Lana Lazebnik

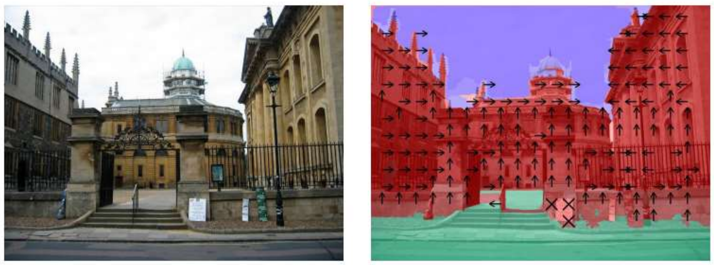
Motion segmentation



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

Regions to surfaces

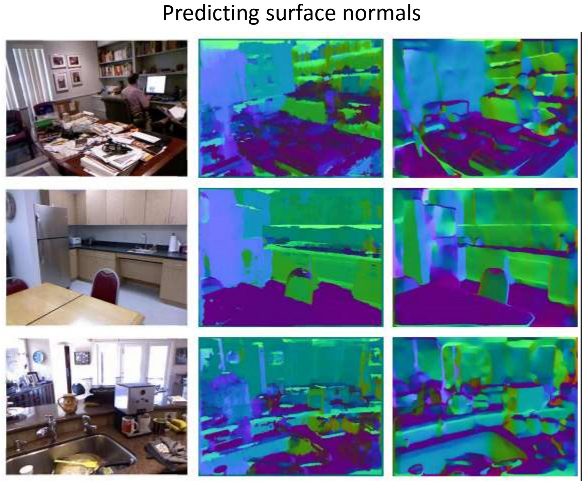
Learn to categorize regions into geometric classes
Combining multiple segmentations



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

Regions to surfaces

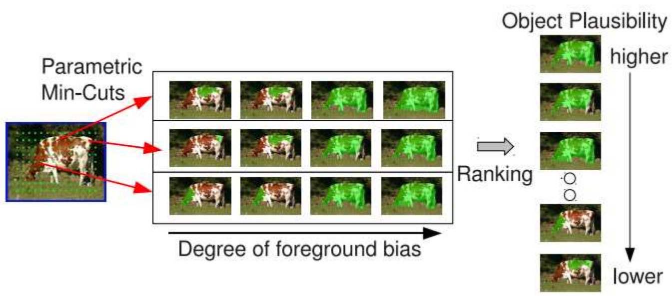
Predicting surface normals



- Ladicky, Zeisl, Pollefeys. Discriminatively Trained Dense Surface Normal Estimation. ECCV 2014

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

Video object segmentation

[Jain & Grauman, Supervoxel-Consistent Foreground Propagation in Video, ECCV 2014]

Kristen Grauman, UTCS

Video object segmentation

Fanyi Xiao and Yong Jae Lee
[Track and Segment: An Iterative Unsupervised Approach for Video Object Proposals]
In *CVPR 2016*