CS9840
Machine Learning in Computer Vision
Olga Veksler

Lecture 4

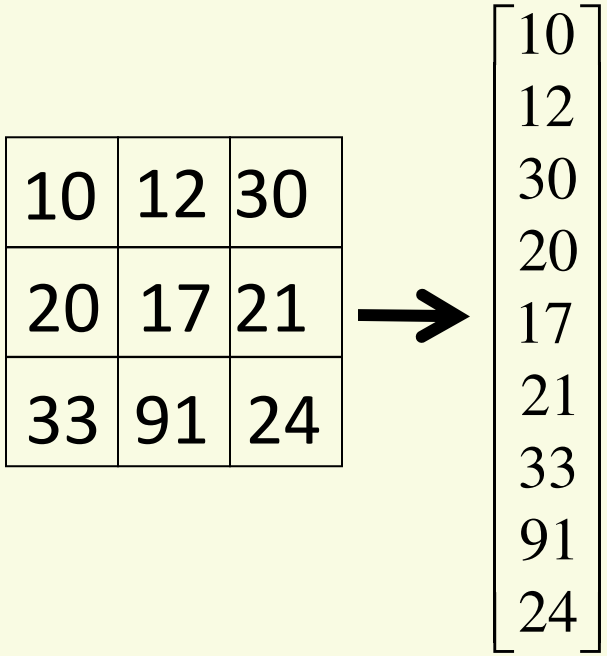
Features for Image Representation

Outline

- How to represent an image?
  - pixel intensity (or color) values
  - global intensity histograms
  - local intensity histograms
  - edge orientation histograms
  - spatial pyramid for histograms

Pixel Representations

- Intensity image
  - pile all values into one vector, say in row order
- Color image
  - Pile each channel values into one vector, say in row order



Slide by Erik Learned-Miller

Pixel Representations

- Small change in image appearance

Slide by Erik Learned-Miller

Pixel Representations

- Leads to a large change in feature vector

difference image

Slide by Erik Learned-Miller

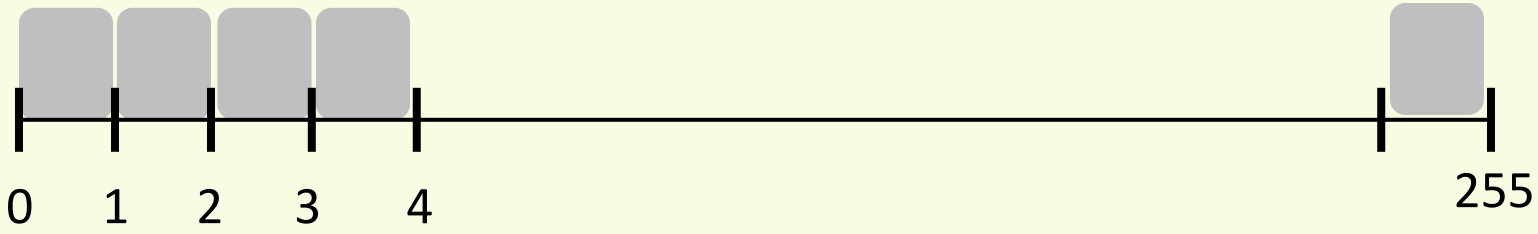
Conclusions

- Pixelwise representations:
  overly sensitive to position
- Nevertheless it has been successfully used in applications
  - eigenfaces, the first successful face detection system

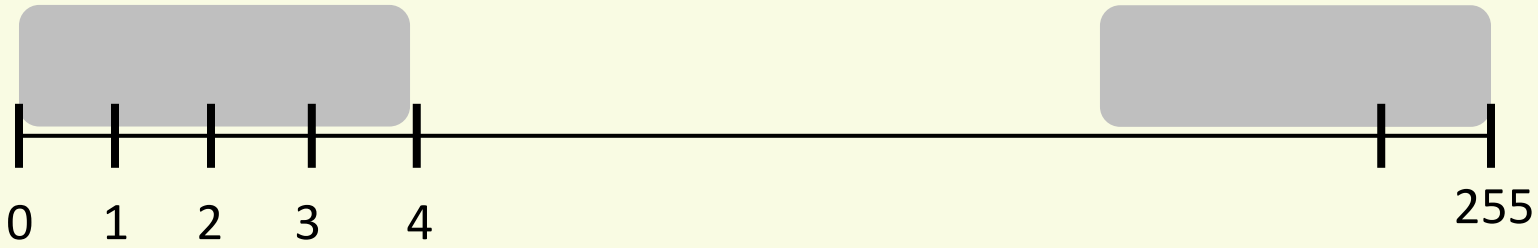
Slide by Erik Learned-Miller

Global Intensity Histogram

- First need to divide intensities in “bins”
- bins size could be at a fine scale:



- or course scale



- This can be looked at as quantization, i.e. larger set of intensities is mapped into a smaller one
  - all values in a bin are treated as one “intensity”. Can display them with the average of their intensities

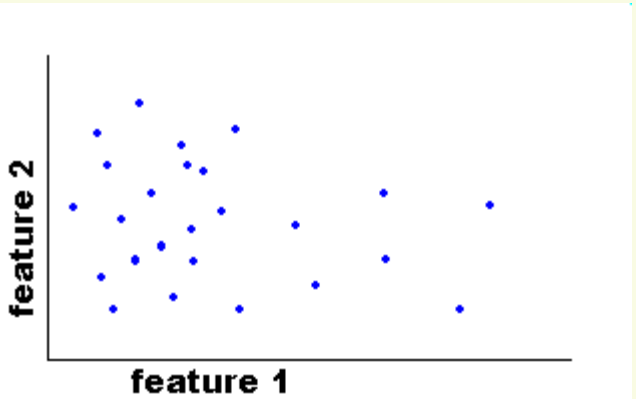
Global Intensity Histogram

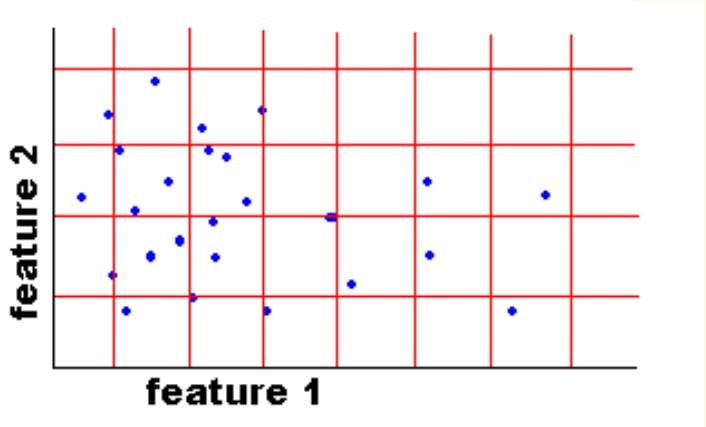
- After bin size is selected, we count data in each bin
- Often use normalized histogram
  - sum up to 1

$$\begin{bmatrix} 10 \\ 20 \\ . \\ . \\ . \\ . \\ . \\ . \\ 240 \end{bmatrix}$$

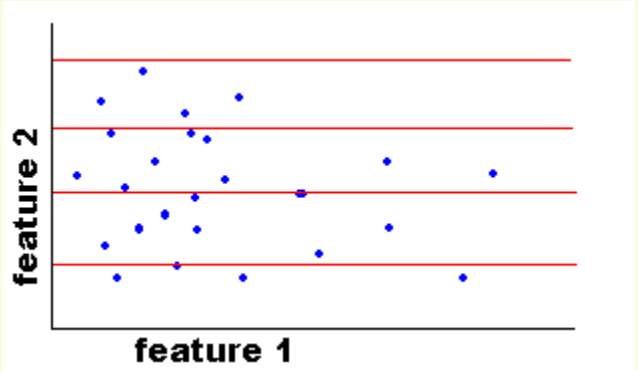
Multi-Dimensional Data, Grid Quantization

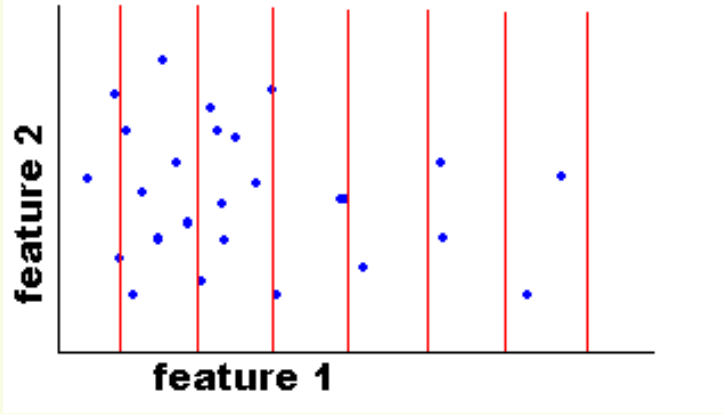
- Suppose we have multi-dimensional data, such as color images
- Different possibilities for histogram





- Joint histogram
  - Requires lots of data
  - Loss of resolution to avoid empty bins





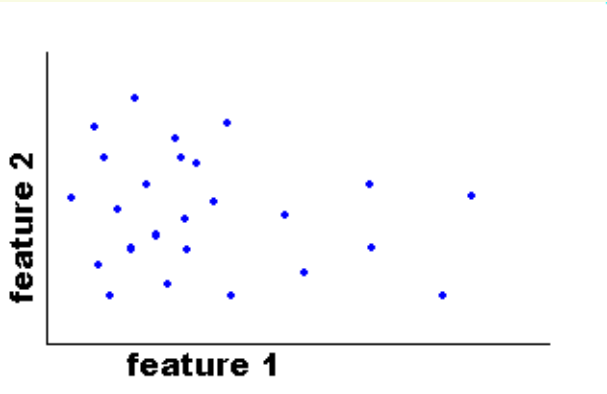
Marginal histogram

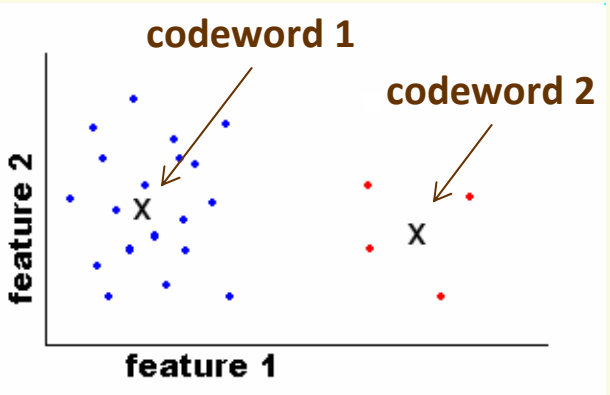
- Ideally, features should be independent
- More data/bin than joint histogram

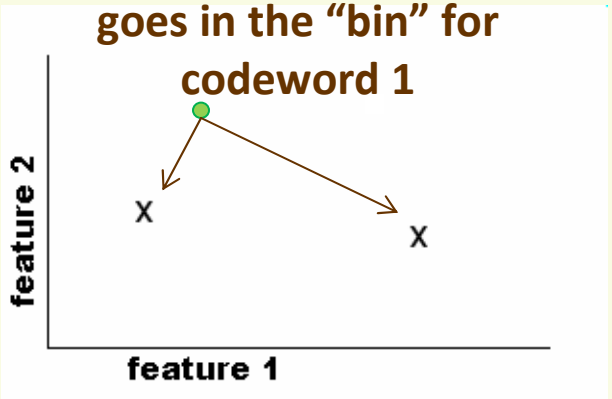
slide credit: Dave Kauchak

Histograms based on Irregular Partitioning

- Can we do irregular partitioning (quantization) ?
- Yes, based on clustering (k-means is often used)





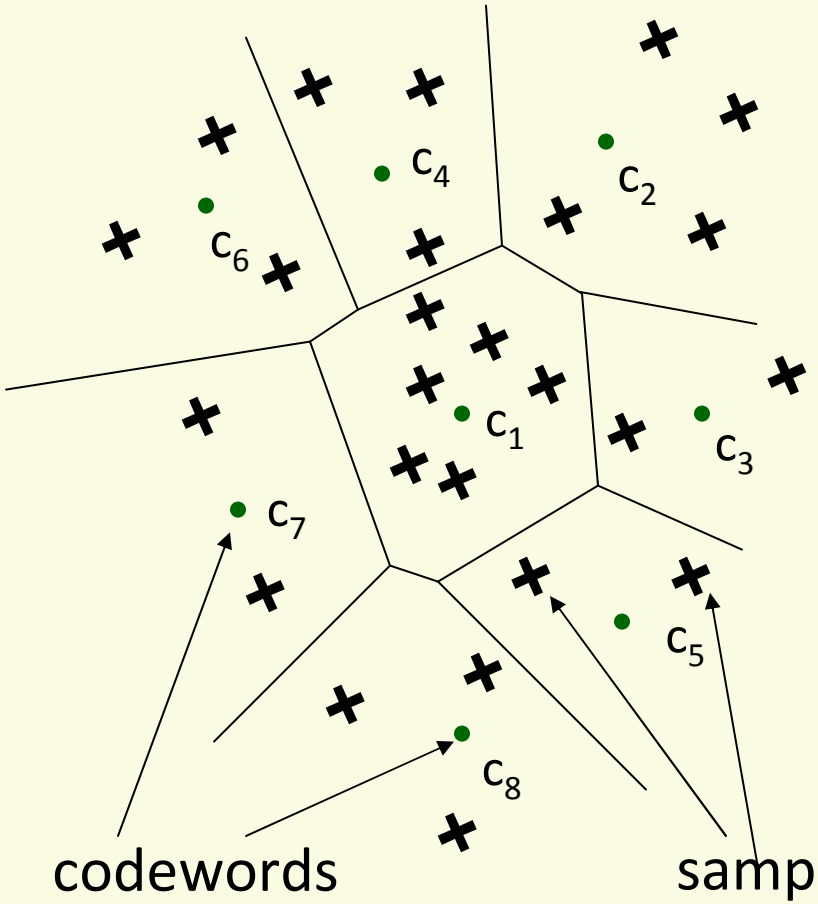


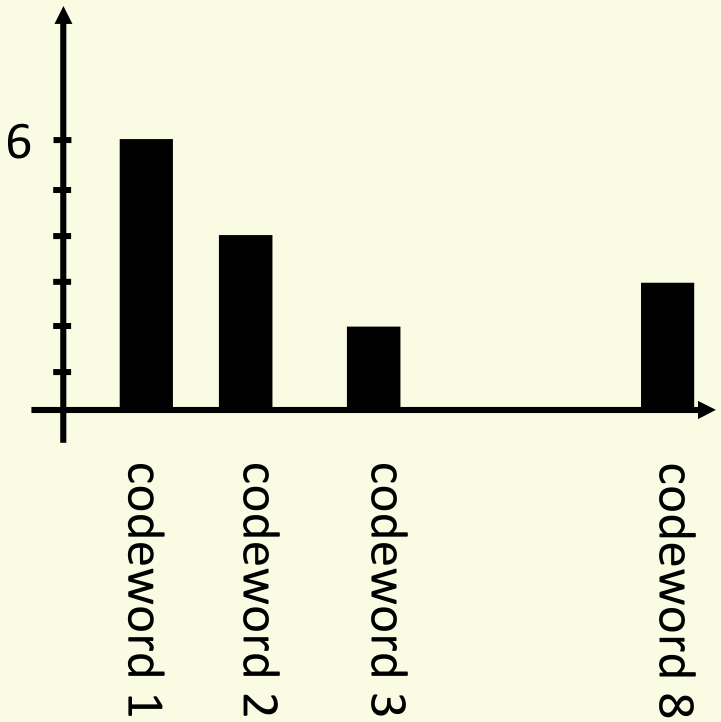
- After clustering, cluster centers (or codewords) stay fixed, these give us “bins” of irregular size
- A sample is mapped to the closest codeword
- Count how many points closest to codeword 1, 2, etc…

Slide Credit: Dave Kauchak

Voronoi Diagram visualization

- Visualization of irregular “bins”





codewords

samples such as

intensity vector, or colors vector, or other image feature vector

Histograms: Implementation issues

- Quantization
  - Grids: fast but applicable only with few dimensions
  - Clustering: slower but can quantize data in higher dimensions
- How many bins?

Few Bins

Need less data

Coarser representation

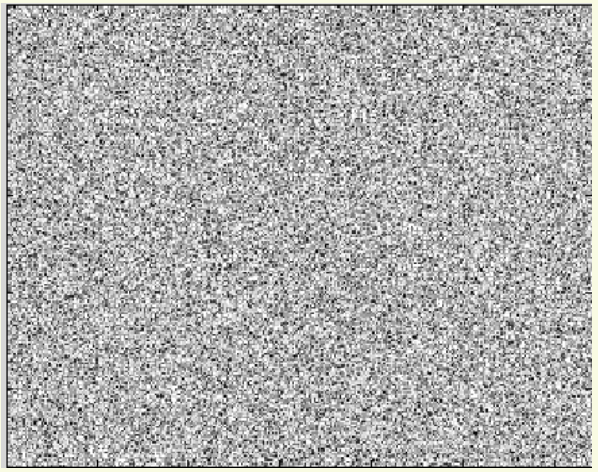
Many Bins

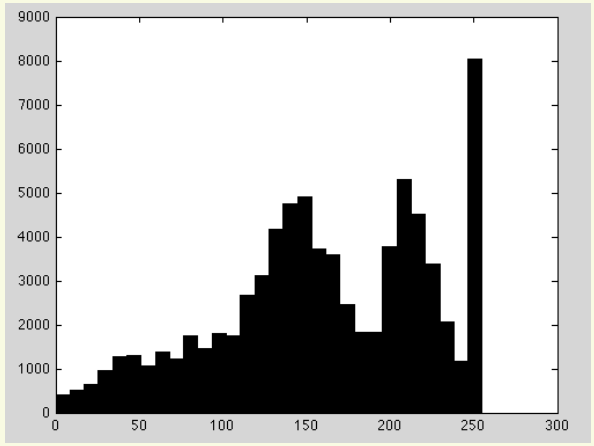
Need more data

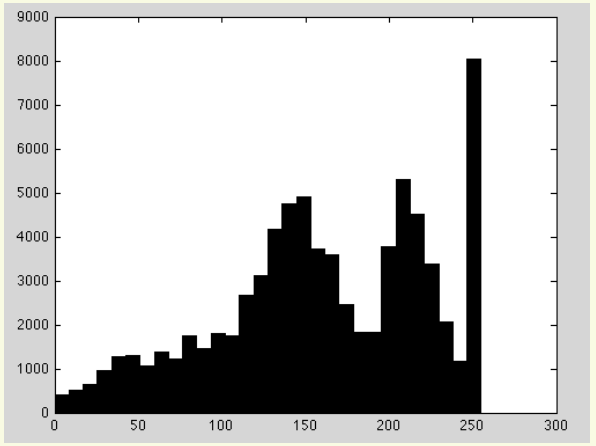
Finer representation

Slide Credit: Derek Hoiem

Problem with Global Histogram



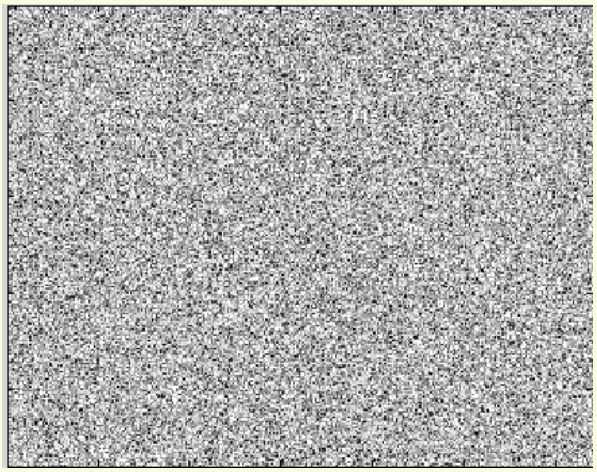




- Identical feature vectors!

Slide by Erik Learned-Miller

Problem with Global Histogram



Have equal histograms!

Slide by Erik Learned-Miller

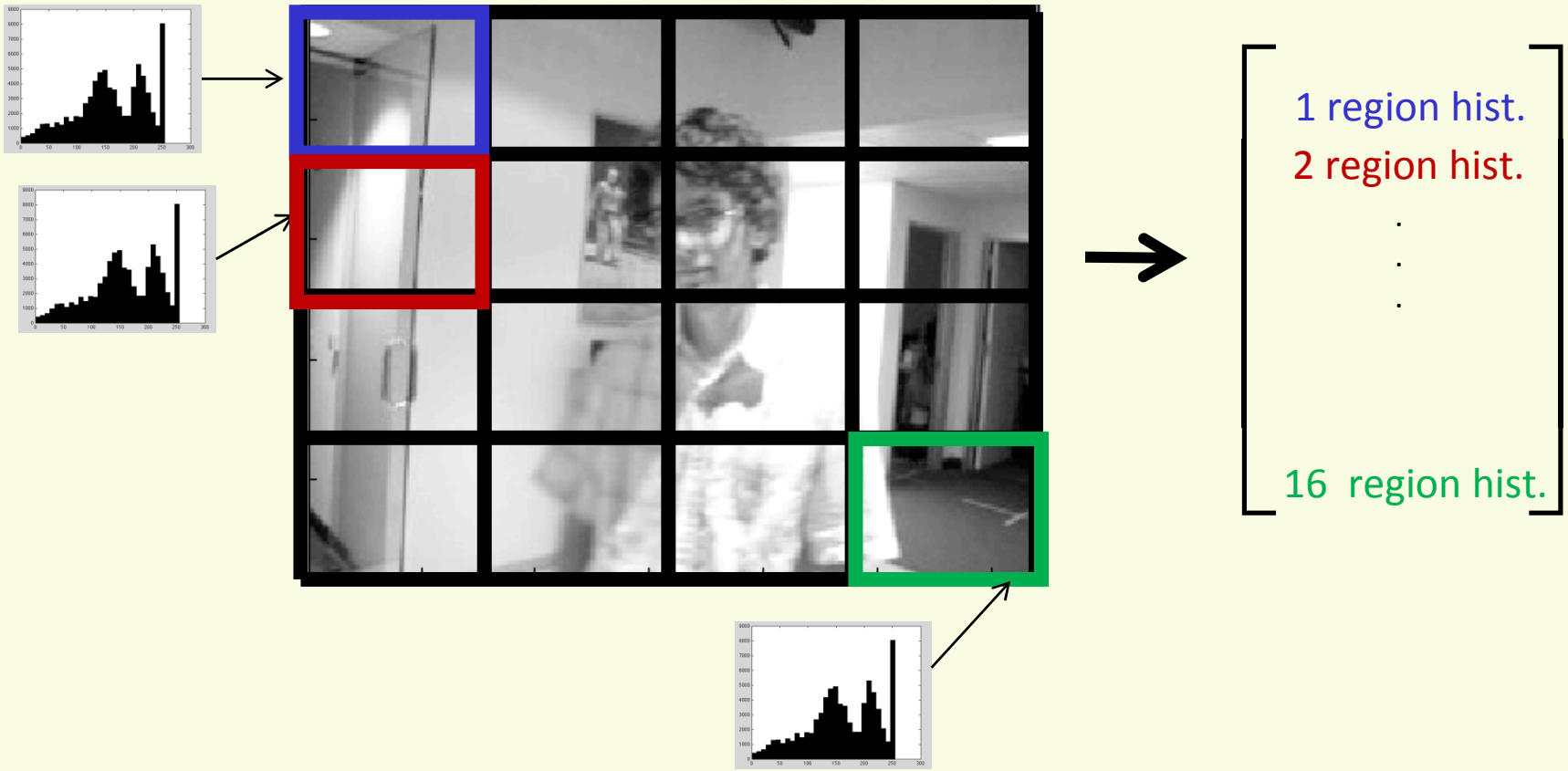
Conclusions

1. Pixel representations:
   overly sensitive to position
2. Global histogram representations:
   under-sensitive to position

Slide by Erik Learned-Miller

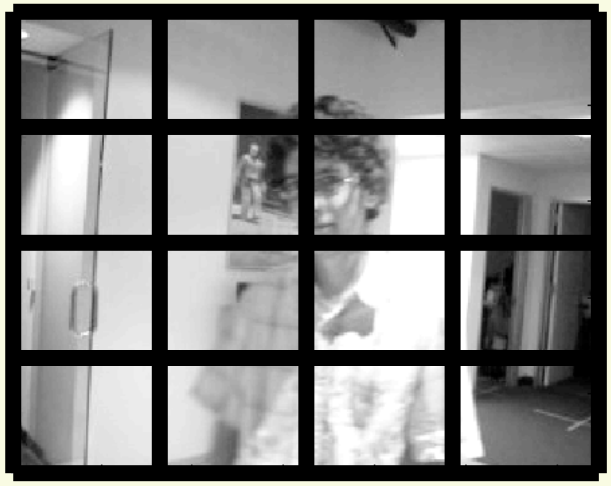
A Compromise: A local histogram

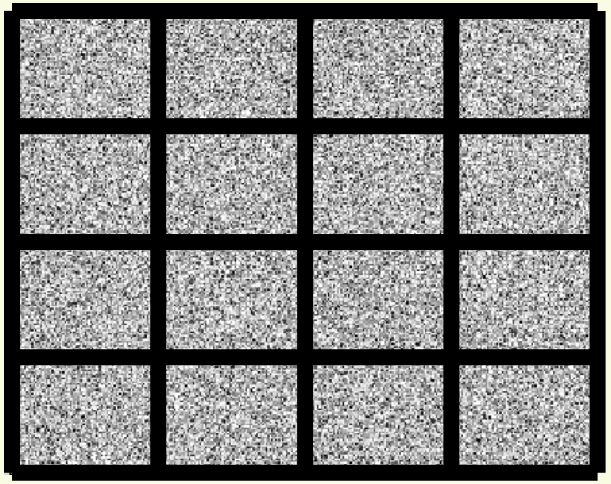
A separate histogram for each region

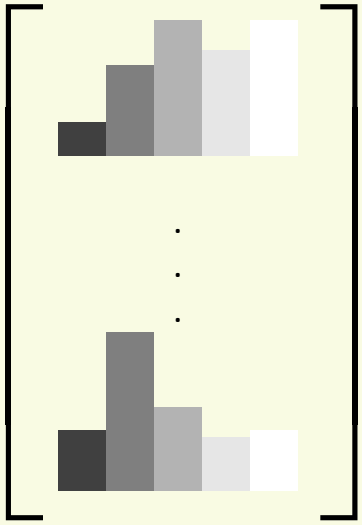


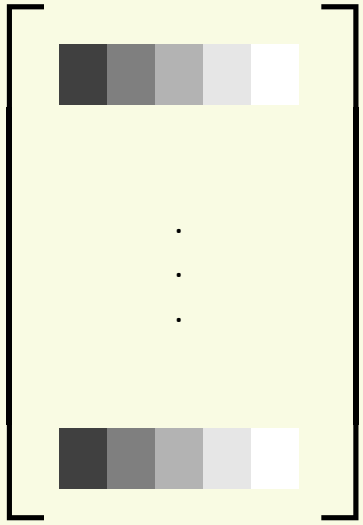
Slide by Erik Learned-Miller

Local Intensity Histogram

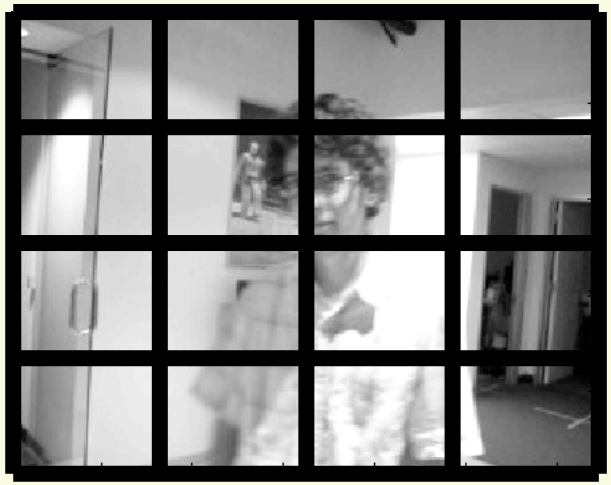


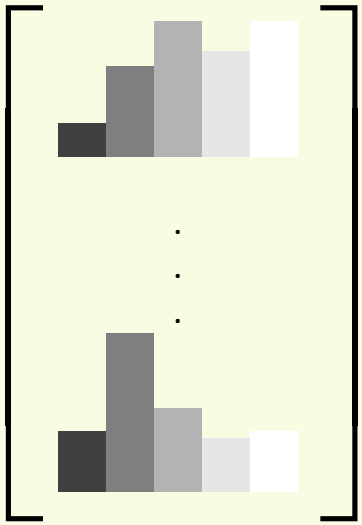


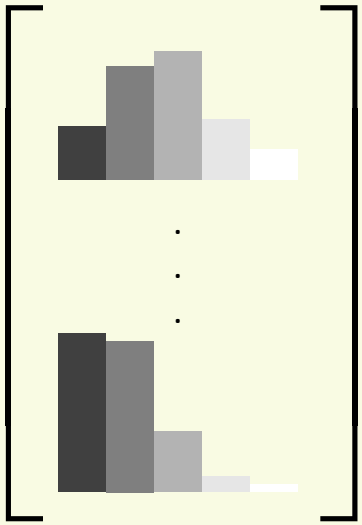




Local Intensity Histogram

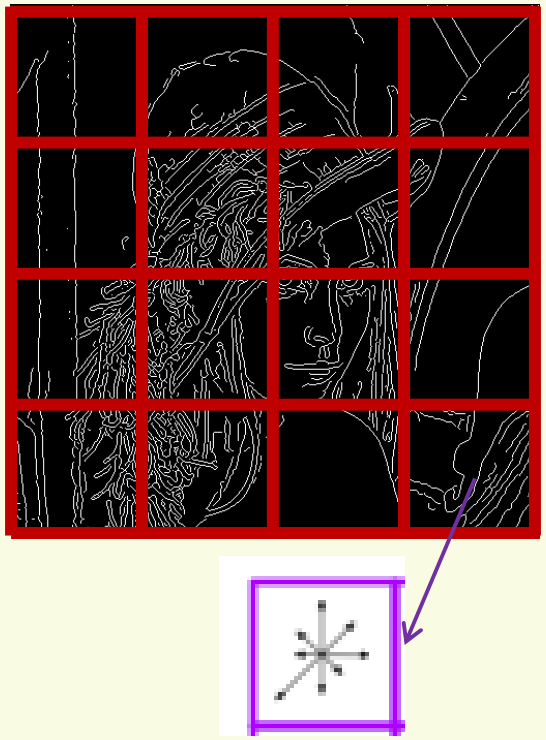


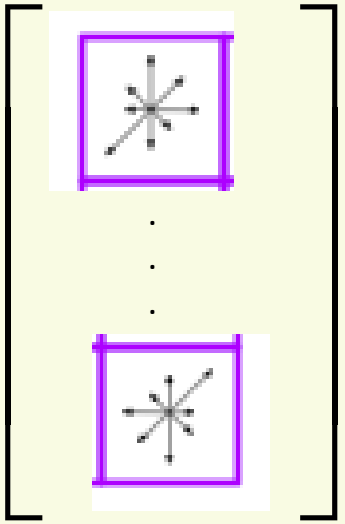




- Intensity histogram is sensitive to lighting changes

Local Edge Orientation Histogram





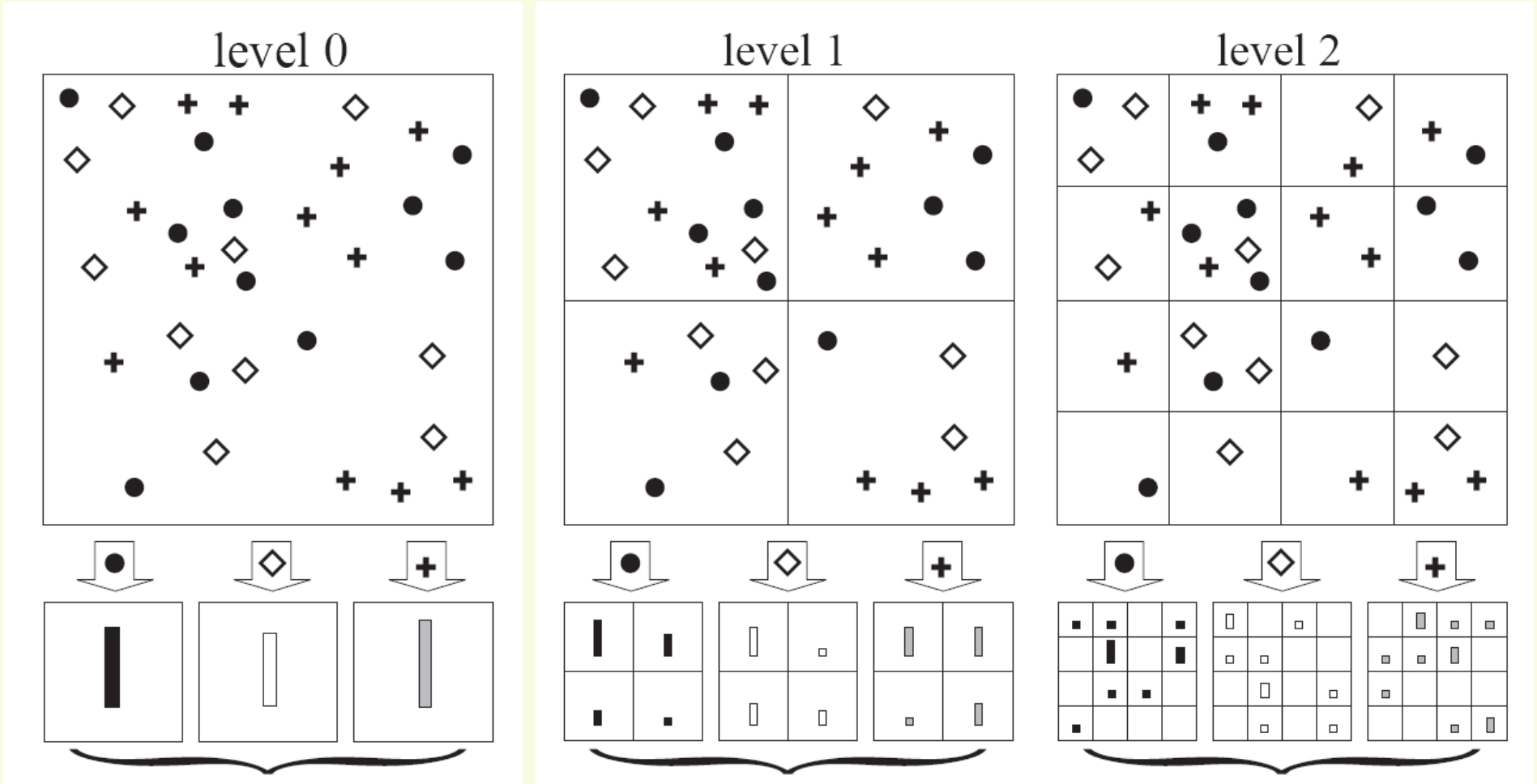
- Edges are not as sensitive to lighting changes
- Compute histogram of edges
  - typically consider only edge orientation
- How do we choose the right box size?

Spatial pyramid

- Use boxes of different sizes!

Slide Credit: Derek Hoiem

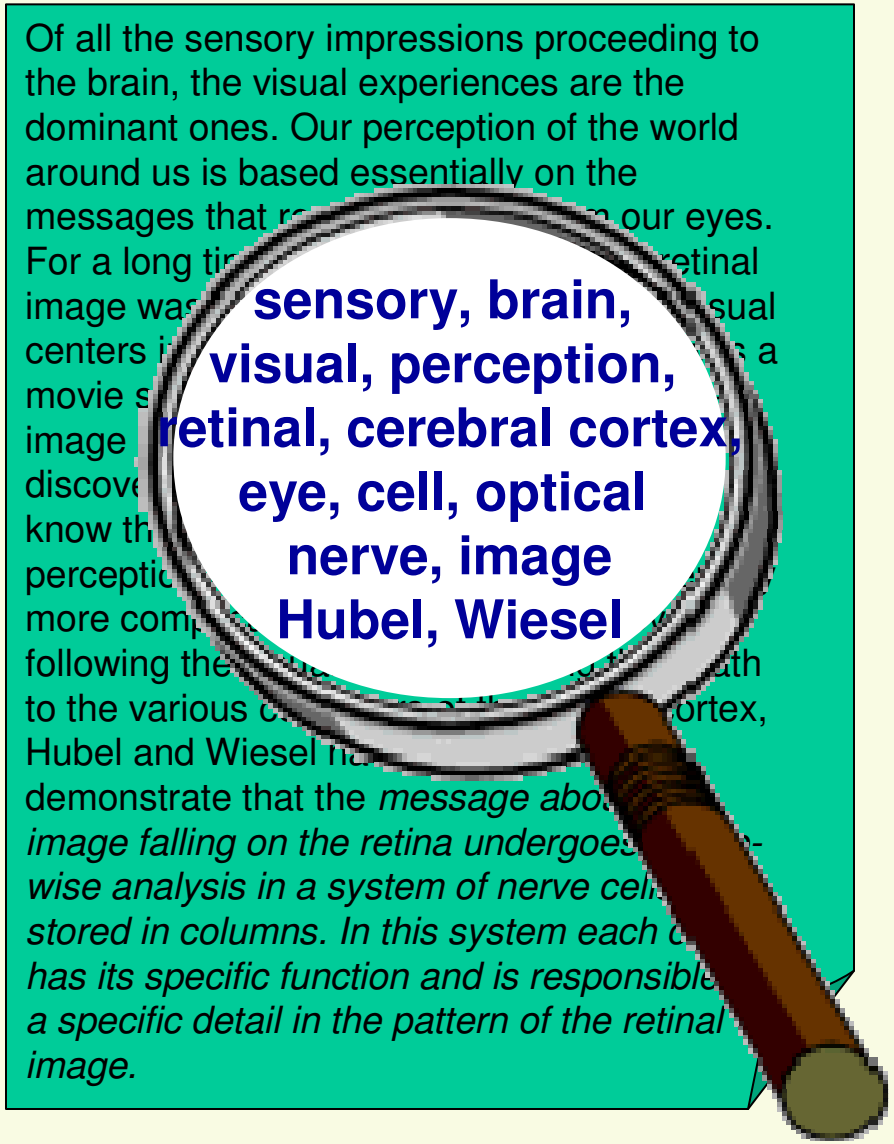
Spatial Pyramid



These get piled up into one feature vector

Slide Credit: Derek Hoiem

Bag of Words (Analogy to documents)

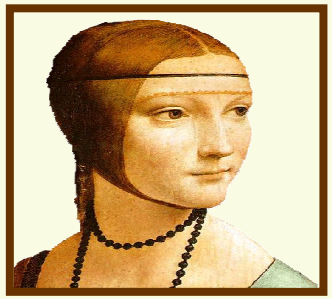
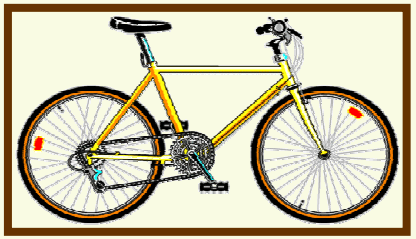


Of all the sensory impressions proceeding to the brain, the visual experiences are the dominant ones. Our perception of the world around us is based essentially on the messages that r… our eyes. For a long ti… retinal image was … sual centers i… a movie s… image … discove… know th… perceptio… more comp… following the … th to the various c… ortex, Hubel and Wiesel ha… demonstrate that the *message abo… image falling on the retina undergoes … wise analysis in a system of nerve cell… stored in columns. In this system each c… has its specific function and is responsible… a specific detail in the pattern of the retinal image.*

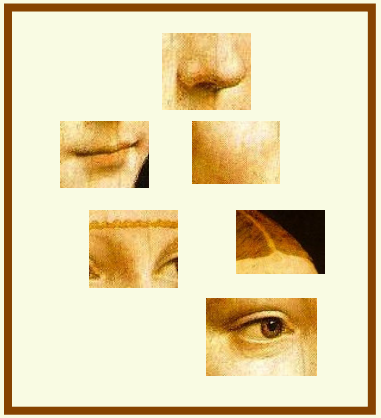
China is forecasting a trade surplus of $90bn (£51bn) to $100bn this year, a threefold increase on 2004's $32bn. The Commerce Ministry said the surplus would be created by a predicted 30% … $750bn, compared wi… $660bn. T… annoy th… China's … deliber… agrees … yuan is … governo… also need… demand so … country. China … e yuan against the do… nd permitted it to trade within a narrow … but the US wants the yuan to be allowed … e freely. However, Beijing has made it c… t it will take its time and tread carefully be… allowing the yuan to rise further in value.

Bag of visual words

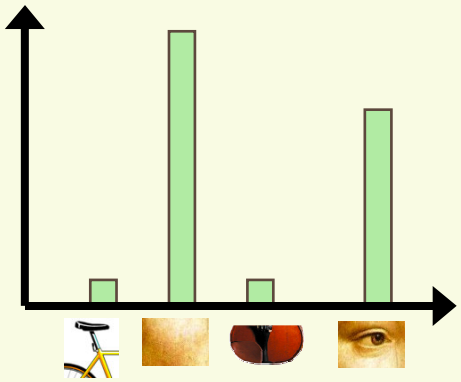
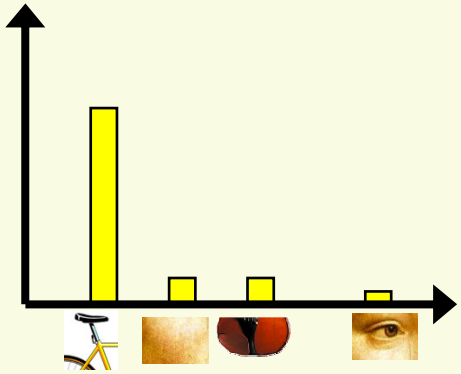
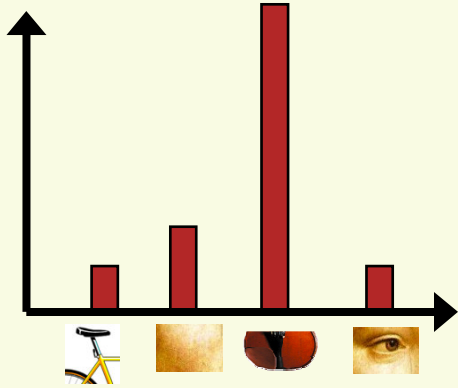
- Training images


- Image patches



- Bow histogram

codewords

Slide by Derek Hoiem

Other Representations

- There are many other ways to represent an image as a feature vector
- Most are based on histogram of
  - texture
  - corner features
  - SIFT features
  - etc.

Right features depend on what you want to know

- Object: 2D shape
  - Local shape info, shading, shadows, texture
- Scene : overall layout
  - linear perspective, gradients
- Material properties: albedo, feel, hardness, …
  - Color, texture
- Motion
  - Optical flow, tracked points

Slide by Derek Hoiem