Lecture 4

Image Representation
Outline

• How to represent an image as a feature vector?
• Basic image features
  • intensity, color, gradients, response to filter(s)
  • dense (at each pixel)
  • sparse (at a subset of locations)
• Representations
  • pixelwise
  • histogram
    • Global vs. Local histograms
    • Spatial pyramids
Basic Image Features

- Given image \( I \), first compute *basic image features or feature responses*
- Then consolidate basic image features into a feature vector \( \mathbf{x} \) that represents image \( I \)
- Simplest basic image feature: intensity of a pixel
  - not enough for most applications
- Other basic image features commonly used:
  - **Color:** 3 values per pixel
  - **Edges:** 1 or 2 values per pixel
  - **Texture:** \( \approx 48 \) values per pixel
Extracting Texture (Texture Responses)

- Texture filter bank:
  - Convolve image with each filter
    - 48 responses per pixel
Kristen Grauman
Kristen Grauman
Extracting Texture

Form a feature vector from the list of responses at each pixel

\[ [r_1, r_2, \ldots, r_{38}] \]
Extracting Texture

[r₁, ..., large, ..., small, ..., r₄₈]

[r₁, ..., small, ..., large, ..., r₄₈]
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
Pixelwise Representation

- Pile basic image feature values into one vector, say row order
- Example: intensity as a basic image feature
  - one value per pixel

\[
\begin{bmatrix}
10 & 12 & 30 \\
20 & 17 & 21 \\
33 & 91 & 24
\end{bmatrix}
\rightarrow
\begin{bmatrix}
10 \\
12 \\
30 \\
20 \\
17 \\
21 \\
33 \\
91 \\
24
\end{bmatrix}
\]
Pixelwise Representation

- Color as a basic image feature
  - three values per pixel
- Pile all color channel into one vector
Pixelwise Representation

- Filter responses as a basic image feature
  - $n$ values per pixel, $n$ is the number of filters
- Pile each filter channel into one vector

```
  5 7 8 3 9 4 3 2
  1 4 2 1 7 1 6 0
```

```plaintext
[5 9 1 7 7 4 4 1 1 8 3 2 6 3 2 1 0]
```
Pixel Representations

- Small change in image appearance
Pixel Representations

- Leads to a large change in feature vector

\[
\begin{bmatrix}
10 & 12 & 30 \\
20 & 17 & 21 \\
33 & 91 & 24
\end{bmatrix} \quad
\begin{bmatrix}
9 & 10 & 12 \\
19 & 20 & 17 \\
32 & 33 & 91
\end{bmatrix}
\]

difference image

\[
[10, 12, 30, 20, 17, 21, 33, 91, 24] \\
[9, 10, 12, 19, 20, 17, 32, 33, 91]
\]

Slide modified from Erik Learned-Miller
Pixel Representations

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

- Think of each intensity value as a “bin”

- Histogram counts the number of values that fall in each bin

- Visual plot:
Global Intensity Histogram

- Insensitive to changes in pixel location
- Often use normalized histogram
  - sums up to 1
Global Intensity Histogram Quantization

- Can quantize intensities (larger bins)

- Histogram: count number of values that fall in each bin

- Quantization
  - helps to improve efficiency
  - groups similar values together (i.e. removes fine distinction)
    - may help for recognition
Multi-Dimensional Histograms

• Basic image features most often multi-dimensional
  • color, texture, optical flow, etc.
• How to build histogram?
• Have to quantize, too sparse without quantization
How to Quantize Multi-Dimensional Data?

1. Joint histogram
   - Need lots of data to avoid empty bins
   - Make bins coarse to simulate lots of data → loose resolution

2. Marginal histogram
   - more data per bin than joint histogram
   - works best for independent features
     - loose correlation information

slide credit: Dave Kauchak
Histograms based on Irregular Partitioning

- Irregular quantization (clustering) gives meaningful bins that adapt to data
  - k-means clustering, etc.

- Cluster centers are called **codewords**
- A sample is identified (assigned to) with the closest codeword
- Build histogram over the codeword
  - count how many samples are closest to codeword 1, codeword 2, etc.
- Need to store only the codewords
Encoding Image $I$ as Feature Vector

- Pre-computed code-words in green
- Extract 2D features from image $I$

- Feature vector that represents image $I$
  - can also normalize it
Clustered Patches

- So far clustered feature responses at each pixel
- Can cluster other things
- Like image patches
  - overlapping or not

![Clustered patches diagram]
Clustered Image Patches
Feature Vector for image $I$

- **Count of codewords**
- **Codewords**

A. Efros
Codewords

- Find codewords on training data, not just one image
- Usually use only a subset of training data for speed

- But not on test data
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 reach us through our eyes. For a long time it was thought that the retinal image was transmitted point by point to visual centers in the brain; the cerebral cortex was a movie screen, so to speak, upon which the image in the eye was projected. Through the discoveries of Hubel and Wiesel we now know that behind the origin of the visual perception in the brain there is a considerably more complicated course of events. By following the visual impulses along their path to the various cell layers of the optical cortex, Hubel and Wiesel have been able to demonstrate that the message about the image falling on the retina undergoes a stepwise analysis in a system of nerve cells stored in columns. In this system each cell has its specific function and is responsible for 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% increase in exports to $750bn, compared with a 18% rise in imports to $660bn. The figures are likely to further annoy the US, which has long argued that China's exports are unfairly helped by a deliberately undervalued yuan. Beijing agrees the surplus is too high, but says the yuan is only one factor. Bank of China governor Zhou Xiaochuan said the country also needed to do more to boost domestic demand so that more goods stayed within the country. China increased the value of the yuan against the dollar by 2.1% in July and permitted it to trade within a narrow band, but the US wants the yuan to be allowed to trade freely. However, Beijing has made it clear that it will take its time and tread carefully before allowing the yuan to rise further in value.

• Inspiration comes from text classification

sensory, brain, visual, perception, retinal, cerebral cortex, eye, cell, optical nerve, image
Hubel, Wiesel

China, trade, surplus, commerce, exports, imports, US, yuan, bank, domestic, foreign, increase, trade, value
Bag of visual words

• Training images

• codewords or visual words

• Bow histogram

codewords
learning

build codewords

codewords dictionary

image representation

Train Classifier

classification

recognition

category decision

A. Efros
Histograms: Implementation issues

• Quantization
  • Grids: fast but applicable only with few dimensions
  • Clustering: slower but can quantize data in higher dimension

• How many bins (clusters)?

<table>
<thead>
<tr>
<th>Few Bins</th>
<th>Many Bins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need less data</td>
<td>Need more data</td>
</tr>
<tr>
<td>Coarser representation</td>
<td>Finer representation</td>
</tr>
<tr>
<td>If too coarse, distinction is lost</td>
<td>If too fine, more distinction than necessary</td>
</tr>
</tbody>
</table>

Slide Credit: Derek Hoiem
Problem with Global Histogram

- Identical feature vectors!
Problem with Global Histogram

Have equal histograms!
Conclusions

1. Pixel representations: 
   *overly sensitive to position*

2. Global histogram representations: 
   *under-sensitive to position*
A Compromise: A local histogram

A separate (normalized) histogram for each region
Local Intensity Histogram

![Image of local intensity histograms]
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!
Spatial Pyramid

These get piled up into one feature vector

Slide Credit: Derek Hoiem
Other Representations

• Many image representation schemes are based on histogram of
  • texture
  • corner features
  • SIFT features
  • etc.

• There are other ways to represent an image as a feature vector