Face Detection and Recognition

Computational Photography
Derek Hoiem, University of Illinois

Lecture by Kevin Karsch

Some slides from Lana Lazebnik, Silvio Savarese, Fei-Fei Li
Administrative stuff

• Final project write-up due Dec 17\textsuperscript{th} 11:59pm
• Presentations Dec 18\textsuperscript{th} 1:30-4:30pm (SC 1214)
• Exams back at end of class
Face detection and recognition

Detection → Recognition → “Sally”
Applications of Face Recognition

- Digital photography
Applications of Face Recognition

- Digital photography
- Surveillance
Applications of Face Recognition

• Digital photography
• Surveillance
• Album organization
Consumer application: iPhoto 2009

http://www.apple.com/ilife/iphoto/
Consumer application: iPhoto 2009

• Can be trained to recognize pets!

What does a face look like?
What does a face look like?
What makes face detection hard?

Expression
What makes face detection hard?

Viewpoint
What makes face detection hard?

Occlusion
What makes face detection and recognition hard?

Coincidental textures
Consumer application: iPhoto 2009

- Things iPhoto thinks are faces
How to find faces anywhere in an image?
Face detection: sliding windows

How to deal with multiple scales?
Face detection
What features?

Exemplars
(Sung Poggio 1994)

Intensity Patterns (with NNs)
(Rowely Baluja Kanade 1996)

Edge (Wavelet) Pyramids
(Schneiderman Kanade 1998)

Haar Filters
(Viola Jones 2000)
How to classify?

• Many ways
  – Neural networks
  – Adaboost
  – SVMs
  – Nearest neighbor
Statistical Template

• Object model = log linear model of parts at fixed positions

\[ +3 +2 -2 -1 -2.5 = -0.5 > 7.5 \]

Non-object

\[ +4 +1 +0.5 +3 +0.5 = 10.5 > 7.5 \]

Object
Training multiple viewpoints

Train new detector for each viewpoint.
Results: faces

Table 1. Face detection with out-of-plane rotation

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>Detection (all faces)</th>
<th>Detection (profiles)</th>
<th>False Detections</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>92.7%</td>
<td>92.8%</td>
<td>700</td>
</tr>
<tr>
<td>1.5</td>
<td>85.5%</td>
<td>86.4%</td>
<td>91</td>
</tr>
<tr>
<td>2.5</td>
<td>75.2%</td>
<td>78.6%</td>
<td>12</td>
</tr>
</tbody>
</table>

208 images with 441 faces, 347 in profile
Face recognition

Detection

Recognition

“Sally”
Face recognition

• Typical scenario: few examples per face, identify or verify test example

• What’s hard: changes in expression, lighting, age, occlusion, viewpoint

• Basic approaches (all nearest neighbor)
  1. Project into a new subspace (or kernel space) (e.g., “Eigenfaces”=PCA)
  2. Measure face features
  3. Make 3d face model, compare shape+appearance (e.g., AAM)
Simple idea

1. Treat pixels as a vector

2. Recognize face by nearest neighbor neighbor

\[ k = \text{argmin}_k \left\| y_k^T - x \right\| \]
The space of all face images

- When viewed as vectors of pixel values, face images are extremely high-dimensional
  - 100x100 image = 10,000 dimensions
  - Slow and lots of storage
- But very few 10,000-dimensional vectors are valid face images
- We want to effectively model the subspace of face images
The space of all face images

- Eigenface idea: construct a low-dimensional linear subspace that best explains the variation in the set of face images
Principle Component Analysis (PCA)
Eigenfaces example

- Training images
- $x_1, \ldots, x_N$
Eigenfaces example

Mean: $\mu$

Top eigenvectors: $u_1, \ldots, u_k$
Visualization of eigenfaces

Principal component (eigenvector) $u_k$

$\mu + 3\sigma_k u_k$

$\mu - 3\sigma_k u_k$
Representation and reconstruction

- Face $\mathbf{x}$ in “face space” coordinates:

$$
\mathbf{x} \rightarrow [\mathbf{u}_1^T (\mathbf{x} - \mu), \ldots, \mathbf{u}_k^T (\mathbf{x} - \mu)] \\
= w_1, \ldots, w_k
$$
Representation and reconstruction

- Face $\mathbf{x}$ in "face space" coordinates:

$$\mathbf{x} \rightarrow [\mathbf{u}_1^T(\mathbf{x} - \mu), \ldots, \mathbf{u}_k^T(\mathbf{x} - \mu)]$$

$$= \mathbf{w}_1, \ldots, \mathbf{w}_k$$

- Reconstruction:

$$\hat{\mathbf{x}} = \mu + w_1\mathbf{u}_1 + w_2\mathbf{u}_2 + w_3\mathbf{u}_3 + w_4\mathbf{u}_4 + \ldots$$
Reconstruction

After computing eigenfaces using 400 face images from ORL face database
Recognition with eigenfaces

Process labeled training images
• Find mean \( \mu \) and covariance matrix \( \Sigma \)
• Find \( k \) principal components (eigenvectors of \( \Sigma \)) \( u_1, \ldots, u_k \)
• Project each training image \( x_i \) onto subspace spanned by principal components:
  \[ (w_{i1}, \ldots, w_{ik}) = (u_1^T(x_i - \mu), \ldots, u_k^T(x_i - \mu)) \]

Given novel image \( x \)
• Project onto subspace:
  \[ (w_1, \ldots, w_k) = (u_1^T(x - \mu), \ldots, u_k^T(x - \mu)) \]
• Optional: check reconstruction error \( x - \hat{x} \) to determine whether image is really a face
• Classify as closest training face in \( k \)-dimensional subspace

PCA

• General dimensionality reduction technique

• Preserves most of variance with a much more compact representation
  – Lower storage requirements (eigenvectors + a few numbers per face)
  – Faster matching

• What are the problems for face recognition?
Limitations

Global appearance method: not robust to misalignment, background variation
Face recognition by humans

Face recognition by humans: 20 results (2005)
Result 1

- Humans can recognize faces in extremely low resolution images.
Result 3

- High-frequency information by itself does not lead to good face recognition performance.
Result 5

- Eyebrows are among the most important for recognition
Result 6

- Both internal and external facial cues are important and they exhibit non-linear interactions.
The important configural relations appear to be independent across the width and height dimensions.
Result 8

- Vertical inversion dramatically reduces recognition performance
Contrast polarity inversion dramatically impairs recognition performance, possibly due to compromised ability to use pigmentation cues.
Result 20

- Human memory for briefly seen faces is rather poor
Fun with Faces

Many slides from Alyosha Efros

Photos From faceresearch.org
The Power of Averaging
The Power of Averaging
8-hour exposure
Averages: Hundreds of images containing a person are averaged to reveal regularities in the intensity patterns across all the images.
How do we average faces?

http://www2.imm.dtu.dk/~aam/datasets/datasets.html
Morphing

image #1

warp

morphing

image #2

warp
Cross-Dissolve vs. Morphing

Average of Appearance Vectors

Average of Shape Vectors

http://www.faceresearch.org/demos/vector

Images from James Hays
Average of multiple Face

1. Warp to mean shape
2. Average pixels

http://www.faceresearch.org/demos/average

### Appearance Vectors vs. Shape Vectors

<table>
<thead>
<tr>
<th>Appearance Vector</th>
<th>Shape Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>200*150 pixels (RGB)</td>
<td>43 coordinates (x,y)</td>
</tr>
</tbody>
</table>

- **Appearance Vector**
  - Requires Annotation
  - Provides alignment!

- **Shape Vector**
  - **Vector of 200*150*3 Dimensions**
  - **Vector of 43*2 Dimensions**

---

**Appearance Vector**

- **Vector of 200*150*3 Dimensions**
- **Requires Annotation**
- **Provides alignment!**

**Shape Vector**

- **Vector of 43*2 Dimensions**
Average Men of the world
Average Women of the world

Central African
Burmes
Cambodian
English
Ethiopian
Filipino
Greek
Indian
Iranian
Irish
Israeli
Italian
Peruvian
Polish
Romanian
Russian
Samoa
South African
Subpopulation means

• Other Examples:
  – Average Kids
  – Happy Males
  – Etc.
  – http://www.faceresearch.org

Average kid

Average happy male

Average female

Average male
Manipulating faces

- How can we make a face look more female/male, young/old, happy/sad, etc.?
- [http://www.faceresearch.org/demos/transform](http://www.faceresearch.org/demos/transform)
Manipulating faces

- We can imagine various meaningful directions.
Face Space

• How to find a set of directions to cover all space?

• We call these directions **Basis**

• If number of basis faces is large enough to span the face subspace:

• Any new face can be represented as a linear combination of basis vectors.

\[ s = \alpha_1 \cdot \text{face}_1 + \alpha_2 \cdot \text{face}_2 + \alpha_3 \cdot \text{face}_3 + \alpha_4 \cdot \text{face}_4 + \ldots = S \cdot a \]
Midterm results + review
Midterm results + review

Mean = 87
Median = 90
Std dev = 7.5