Face Recognition

CSE 576
Face recognition: once you’ve detected and cropped a face, try to recognize it

Detection

Recognition

“Sally”
Face recognition: overview

• 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
  2. Measure face features
Typical face recognition scenarios

• Verification: a person is claiming a particular identity; verify whether that is true
  – E.g., security

• Closed-world identification: assign a face to one person from among a known set

• General identification: assign a face to a known person or to “unknown”
What makes face recognition hard?

Expression
What makes face recognition hard?

Lighting
What makes face recognition hard?

Occlusion
What makes face recognition hard?

Viewpoint
Simple idea for face recognition

1. Treat face image as a vector of intensities

2. Recognize face by nearest neighbor in database

\[ k = \arg\min_k \|y_k - x\| \]
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

- Idea: construct a low-dimensional linear subspace that best explains the variation in the set of face images
Consider the variation along direction $\mathbf{v}$ among all of the orange points:

$$\text{var}(\mathbf{v}) = \sum_{\text{orange point } \mathbf{x}} \| (\mathbf{x} - \bar{\mathbf{x}})^T \cdot \mathbf{v} \|^2$$

What unit vector $\mathbf{v}$ minimizes $\text{var}$?

$$\mathbf{v}_2 = \min_{\mathbf{v}} \{ \text{var}(\mathbf{v}) \}$$

What unit vector $\mathbf{v}$ maximizes $\text{var}$?

$$\mathbf{v}_1 = \max_{\mathbf{v}} \{ \text{var}(\mathbf{v}) \}$$

Solution: $\mathbf{v}_1$ is eigenvector of $A$ with largest eigenvalue

$\mathbf{v}_2$ is eigenvector of $A$ with smallest eigenvalue
Principal component analysis (PCA)

• Suppose each data point is N-dimensional
  – Same procedure applies:

  \[
  \text{var}(v) = \sum_x \| (x - \bar{x})^T \cdot v \|
  = v^T A v \quad \text{where} \quad A = \sum_x (x - \bar{x})(x - \bar{x})^T
  \]

  – The eigenvectors of A define a new coordinate system
    • eigenvector with largest eigenvalue captures the most variation among training vectors x
    • eigenvector with smallest eigenvalue has least variation
  – We can compress the data by only using the top few eigenvectors
    • corresponds to choosing a “linear subspace”
      – represent points on a line, plane, or “hyper-plane”
    • these eigenvectors are known as the principal components
The space of faces

• An image is a point in a high dimensional space
  – An N x M image is a point in $\mathbb{R}^{NM}$
  – We can define vectors in this space as we did in the 2D case
Dimensionality reduction

• The set of faces is a “subspace” of the set of images
  – Suppose it is K dimensional
  – We can find the best subspace using PCA
  – This is like fitting a “hyper-plane” to the set of faces
    • spanned by vectors $v_1, v_2, ..., v_K$
    • any face $x \approx \bar{x} + a_1v_1 + a_2v_2 + \ldots + a_kv_k$
Eigenfaces

• PCA extracts the eigenvectors of $A$
  - Gives a set of vectors $v_1, v_2, v_3, ...$
  - Each one of these vectors is a direction in face space
    • what do these look like?
Visualization of eigenfaces

Principal component (eigenvector) $u_k$

$\mu + 3\sigma_k u_k$

$\mu - 3\sigma_k u_k$
Projecting onto the eigenfaces

- The eigenfaces $v_1, ..., v_K$ span the space of faces
  - A face is converted to eigenface coordinates by
    \[
    \mathbf{x} \rightarrow \left( (\mathbf{x} - \bar{x}) \cdot v_1, (\mathbf{x} - \bar{x}) \cdot v_2, \ldots, (\mathbf{x} - \bar{x}) \cdot v_K \right)
    \]
    
    \[
    \mathbf{x} \approx \bar{x} + a_1 v_1 + a_2 v_2 + \ldots + a_K v_K
    \]
Recognition with eigenfaces

• Algorithm

1. Process the image database (set of images with labels)
   • Run PCA—compute eigenfaces
   • Calculate the $K$ coefficients for each image

2. Given a new image (to be recognized) $x$, calculate $K$ coefficients
   \[ x \rightarrow (a_1, a_2, \ldots, a_K) \]

3. Detect if $x$ is a face
   \[ \| x - (\bar{x} + a_1 v_1 + a_2 v_2 + \ldots + a_K v_K) \| < \text{threshold} \]

4. If it is a face, who is it?
   • Find closest labeled face in database
   • nearest-neighbor in $K$-dimensional space
Choosing the dimension $K$

- How many eigenfaces to use?
- Look at the decay of the eigenvalues
  - the eigenvalue tells you the amount of variance “in the direction” of that eigenface
  - ignore eigenfaces with low variance
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
Enhancing gender

Changing age

• Face becomes “rounder” and “more textured” and “grayer”

• Original

• Color

Which face is more attractive?

http://www.beautycheck.de
Which face is more attractive?

right
Which face is more attractive?

0.5(attractive + average)  attractive
Which face is more attractive?

http://www.beautycheck.de
Which face is more attractive?

Right
Which face is more attractive?

0.5(adult+child)  adult
Limitations

- The direction of maximum variance is not always good for classification
A more discriminative subspace: FLD

• Fisher Linear Discriminants $\rightarrow$ “Fisher Faces”

• PCA preserves maximum variance

• FLD preserves discrimination
  – Find projection that maximizes scatter between classes and minimizes scatter within classes

Reference: Eigenfaces vs. Fisherfaces, Belheumer et al., PAMI 1997
Illustration of the Projection

- Using two classes as example:

Poor Projection

Good
Comparing with PCA
Variables

- N Sample images: \(\{x_1, \ldots, x_N\}\)
- c classes: \(\{\chi_1, \ldots, \chi_c\}\)
- Average of each class: \(\mu_i = \frac{1}{N_i} \sum_{x_k \in \chi_i} x_k\)
- Average of all data: \(\mu = \frac{1}{N} \sum_{k=1}^{N} x_k\)
Scatter Matrices

- Scatter of class $i$: $S_i = \sum_{x_k \in \chi_i} (x_k - \mu_i)(x_k - \mu_i)^T$

- Within class scatter: $S_W = \sum_{i=1}^{c} S_i$

- Between class scatter: $S_B = \sum_{i=1}^{c} N_i (\mu_i - \mu)(\mu_i - \mu)^T$
Illustration

Within class scatter

\[ S_W = S_1 + S_2 \]

Between class scatter
Mathematical Formulation

• After projection
  – Between class scatter
  \[ y_k = W^T x_k \]
  – Within class scatter
  \[ \tilde{S}_B = W^T S_B W \]
  \[ \tilde{S}_W = W^T S_W W \]

• Objective
  \[ W_{opt} = \arg \max_w \left| \frac{\tilde{S}_B}{\tilde{S}_W} \right| = \arg \max_w \left| \frac{W^T S_B W}{W^T S_W W} \right| \]

• Solution: Generalized Eigenvectors
  \[ S_B w_i = \lambda_i S_W w_i \quad i = 1, \ldots, m \]

• Rank of \( W_{opt} \) is limited
  – Rank\((S_B) \) <= |C|-1
  – Rank\((S_W) \) <= N-C
Illustration

$S_W = S_1 + S_2$
Recognition with FLD

• Use PCA to reduce dimensions to N-C
  
  \[ W_{pca} = \text{pca}(X) \]

• Compute within-class and between-class scatter matrices for PCA coefficients

  \[ S_i = \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T \]
  \[ S_w = \sum_{i=1}^{c} S_i \]
  \[ S_B = \sum_{i=1}^{c} N_i (\mu_i - \mu)(\mu_i - \mu)^T \]

• Solve generalized eigenvector problem

  \[ W_{fld} = \arg \max_w \frac{|W^T S_B W|}{|W^T S_w W|} \]
  \[ S_B w_i = \lambda_i S_w w_i \quad i = 1, \ldots, m \]

• Project to FLD subspace (c-1 dimensions)

  \[ \hat{x} = W_{opt}^T x \]

• Classify by nearest neighbor

Note: \( x \) in step 2 refers to PCA coeff; \( x \) in step 4 refers to original data
Results: Eigenface vs. Fisherface

• Input: 160 images of 16 people
• Train: 159 images
• Test: 1 image

• Variation in Facial Expression, Eyewear, and Lighting

Reference: Eigenfaces vs. Fisherfaces, Belheumer et al., PAMI 1997
Eigenfaces vs. Fisherfaces

Reference: Eigenfaces vs. Fisherfaces, Belheumer et al., PAMI 1997
Large scale comparison of methods

- FRVT 2006 Report
FVRT Challenge

- Frontal faces
  - FVRT2006 evaluation

False Rejection Rate at False Acceptance Rate = 0.001
FVRT Challenge

• Frontal faces
  – FVRT2006 evaluation: controlled illumination
FVRT Challenge

- Frontal faces
  - FVRT2006 evaluation: uncontrolled illumination
FVRT Challenge

• Frontal faces
  – FVRT2006 evaluation: computers win!
Face recognition by humans

Face recognition by humans: 20 results (2005)
Humans can recognize faces in extremely low resolution images.
High-frequency information by itself does not lead to good face recognition performance.
Result 4: Facial features are processed holistically

Figure 4. Try to name the famous faces depicted in the two halves of the left image. Now try the right image. Subjects find it much more difficult to perform this task when the halves are aligned (left) compared to misaligned halves (right), presumably because holistic processing interacts (and in this case, interferes) with feature-based processing.

Result 5: Of the different facial features, eyebrows are amongst the most important for recognition. Not all facial features are created equal in terms of their role in helping identify a face. However, one facial feature has, surprisingly, received little attention from researchers in this domain – the eyebrows. Sadri et al (2003) have presented striking new evidence suggesting that the eyebrows might not only be important features, but that they might well be the most important, eclipsing even the eyes. These researchers digitally erased the eyebrows from a set of 50 celebrity face images (figure 5). Subjects were shown these images individually and asked to name them. Subsequently, they were asked to recognize the original set of (unaltered) images. Performance was recorded as the proportion of faces a subject was able to recognize. Performance with the images lacking eyebrows was significantly worse relative to that with the originals, and even with the images lacking eyes. These results suggest that the eyebrows may contribute in an important way to the representations underlying identity assessments.

How might one reason plausibly explain the perceptual significance of eyebrows in face recognition? There are several possibilities. First, eyebrows appear to be very important for conveying emotions and other nonverbal signals. Since the visual system may already be biased to attend to the eyebrows in order to detect and interpret such signals, it may be that this bias also extends to the task of facial identification. Second, for a number of reasons, eyebrows may serve as a very “stable” facial feature. Because they tend to be
Eyebrows are among the most important for recognition
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.
Vertical inversion dramatically reduces recognition performance
Contrast polarity inversion dramatically impairs recognition performance, possibly due to compromised ability to use pigmentation cues.
Motion of faces appears to facilitate subsequent recognition
The visual system starts with a rudimentary preference for face-like patterns.
Result 17: Vision progresses from piecemeal to holistic

<table>
<thead>
<tr>
<th>Age</th>
<th>Faces</th>
<th></th>
<th></th>
<th>Houses</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Upright</td>
<td>Inverted</td>
<td></td>
<td>Upright</td>
<td>Inverted</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>69</td>
<td>64</td>
<td></td>
<td>71</td>
<td>58*†</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>81</td>
<td>67</td>
<td></td>
<td>74</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>89</td>
<td>68‡</td>
<td></td>
<td>73</td>
<td>77</td>
<td></td>
</tr>
</tbody>
</table>

Note: * and † indicate significant differences.
Human memory for briefly seen faces is rather poor.
Things to remember

• PCA is a generally useful dimensionality reduction technique
  – But not ideal for discrimination

• FLD better for discrimination, though only ideal under Gaussian data assumptions

• Computer face recognition works very well under controlled environments – still room for improvement in general conditions