Face Recognition

Haibin Ling

http://www.dabi.temple.edu/~hbling/Teaching/13F_5543/index.html

Many slides revised from K. Grosse, R. Fergus, S. Lazebnik

Preface

- Face recognition
  - Given a test face and a set of reference faces in a database find the N closest reference faces to the test one.
- Face authentication
  - Given a test face and a reference one, decide if the test face is identical to the reference one.

Motivation

- Application Demands
  - Nonintrusive identification
  - Nonintrusive verification
  - Nonintrusive access control
  - Identification for law enforcement

Challenges in face recognition

- Many variations
  - Pose variation
  - Illumination conditions
  - Scale variability
  - Age difference
  - Expression
- Varied image conditions
  - Occlusion
  - Low resolution
  - Noise

Outline

- Holistic face recognition, intensity based
  - Eigenfaces
- Modeling texture and geometry
  - Elastic Bunch Graph Matching
- Shape and appearance
  - Active Appearance models
Principal Component Analysis

- Given: N data points $x_1, \ldots, x_N$ in $\mathbb{R}^d$
- We want to find a new set of features that are linear combinations of original ones:
  $$ u(x_i) = u^T(x_i - \mu) $$
  ($\mu$: mean of data points)
- What unit vector $u$ in $\mathbb{R}^d$ captures the most variance of the data?

Eigenfaces: Key idea

- Assume that most face images lie on a low-dimensional subspace determined by the first $k$ ($k < d$) directions of maximum variance
- Use PCA to determine the vectors $u_1, \ldots, u_k$ that span that subspace:
  $$ x = \mu + w_1 u_1 + w_2 u_2 + \ldots + w_k u_k $$
- Represent each face using its "face space" coordinates $(w_1, \ldots, w_k)$
- Perform nearest-neighbor recognition in "face space"

Eigenface examples

- Training images $x_1, \ldots, x_N$

Eigenface example

- Top eigenvectors: $u_1, \ldots, u_k$

Mean: $\mu$

Eigenfaces example

- Face $x$ in "face space" coordinates:
  $$ x \rightarrow [u_1^T(x - \mu), \ldots, u_k^T(x - \mu)] = [w_1, \ldots, w_k] $$
Eigenfaces example

- Face $x$ in "face space" coordinates:
  \[ x \rightarrow \begin{bmatrix} \mathbf{u}_1^T (x - \mu) \\ \vdots \\ \mathbf{u}_k^T (x - \mu) \end{bmatrix} = w_1, \ldots, w_k \]

Reconstruction:

\[ \hat{x} = \mu + w_1 \mathbf{u}_1 + w_2 \mathbf{u}_2 + \ldots \]

Summary: Recognition with eigenfaces

- Process labeled training images:
  - Find mean $\mu$ and covariance matrix $\Sigma$
  - Find $k$ principal components (eigenvectors of $\Sigma$) $\mathbf{u}_1, \ldots, \mathbf{u}_k$
  - Project each training image $x_i$ onto subspace spanned by principal components:
    \[ (w_{i1}, \ldots, w_{ik}) = (\mathbf{u}_1^T (x_i - \mu), \ldots, \mathbf{u}_k^T (x_i - \mu)) \]

- Given novel image $x$:
  - Project onto subspace:
    \[ (w_{1}, \ldots, w_{k}) = (\mathbf{u}_1^T (x - \mu), \ldots, \mathbf{u}_k^T (x - \mu)) \]
  - Optional: check reconstruction error $x - \hat{x}$ to determine whether image is really a face
  - Classify as closest training face in $k$-dimensional subspace

Limitations

- Global appearance method: not robust to

Limitations

- PCA assumes that the data has a Gaussian distribution (mean $\mu$, covariance matrix $\Sigma$)

Outline

- Holistic face recognition, intensity based
  - Eigenfaces

- Shape and appearance
  - Active Appearance models

- Modeling texture and geometry
  - Elastic Bunch Graph Matching

Other Component Analysis

- Is principle component the right one?
  - Direction of maximum variance good for classification?

- More subspace methods:
  - Fisherfaces (LDA, Belhumeur et al. 1997)
  - Independent Component Analysis (ICA, Bartlett et al. 2002)
  - Nonlinear embedding
    - Laplacian face (LPP, He et al. 2005)
Essence of the Idea: Recognition by Synthesis

- Explain a new example in terms of the model parameters

So what's a model

Model

"Shape"

"texture"

Slide: Dhruv Batra

Active Shape Models

training set

Profile

Half Profile

Frontal

Set of Points

\[ x = (x_1, \ldots, x_n, y_1, \ldots, y_n)^T \]

Provides alignment!

Slide: Dhruv Batra

Shape Vector

The Morphable Face Model

The structure of a face

- Shape vector \( S = (x_1, y_1, x_2, \ldots, y_n)^T \), containing the \((x, y)\) coordinates of vertices of a face.
- Appearance vector \( T = (R_1, G_1, B_1, R_2, \ldots, G_n, B_n)^T \), containing the color values of the mean-warped face image.

The Morphable face model

- Again, assuming that we have \( m \) such vector pairs in full correspondence, we can form new shapes \( S_{\text{model}} \) and new appearances \( T_{\text{model}} \) as:

\[
S_{\text{model}} = \sum_{i=1}^{m+1} \alpha_i S_i, \quad T_{\text{model}} = \sum_{i=1}^{m+1} \beta_i T_i
\]

- If number of basis faces \( m \) is large enough to span the face subspace then:
  - Any new face can be represented as a pair of vectors \((\alpha_1, \alpha_2, \ldots, \alpha_m)\) and \((\beta_1, \beta_2, \ldots, \beta_m)\)!

Playing with the Parameters

First two modes of shape variation

First two modes of gray level variation

First four modes of appearance variation

Overview

- Holistic face recognition, intensity based
  - Eigenfaces
- Shape and appearance
  - Active Appearance models
- Modeling texture and geometry
  - Elastic Bunch Graph Matching

EBGM Overview

- Human faces share a similar topological structure
  - Labeled graph as basic object representation
    - Nodes positioned at fiducial points (eyes, nose...)
    - Jets at each node
    - Edges labeled with distance information
  - Stored model graph matched to new images
    - Image graph (can become model graph)
  - Model graphs easily translated, scaled, orientated

Gabor wavelets

- Shape of plane waves restricted by a Gaussian envelope function
  - Hence good results in practice
  - Biologically motivated

  **Pro:**
  - Invariant to changes in brightness
  - Robust against translation or distortion

  **Con:**
  - Dependent on the background of the image
Gabor wavelets

Family of Gabor kernels

\[ \psi_j(x) = \frac{\partial^2}{\partial x^2} \exp \left( -\frac{x^2}{2\sigma^2} \right) \left[ \exp(i\mathbf{k}_j \cdot \mathbf{x}) - \exp \left( -\frac{\mathbf{k}_j^2}{2\sigma^2} \right) \right] \]

In the shape of plane waves with wave vector \( \mathbf{k}_j \)

restricted by a Gaussian envelope function.

5 different frequencies \( f = 0, \ldots, 4 \)
8 orientations \( \varphi = 0, \pi/4, \ldots, 7\pi/4 \)

\( \mathbf{k}_j = \left( \frac{k_j \sin \varphi_j}{\sqrt{2}}, \frac{k_j \cos \varphi_j}{\sqrt{2}} \right), \quad k_j = 2^{-1/8} f_j, \quad \varphi_j = \frac{j\pi}{8} \)

Width \( \sigma/k \) of Gaussian controlled by \( \sigma = 2\pi \)

Family of kernels is self-similar and generated from one mother wavelet by dilation and rotation!

Jets

- Wavelets for different frequencies and orientation
- Jet describes a small patch of grey values
- Defined as the set of complex coefficients
  \[ J = \{ J_i \} \]
  for a given pixel

Image graph

- Image Graph G: N nodes, E edges
- Labeling of nodes:
  Jets \( J_n \) at positions \( x_n, n = 1, \ldots, N \)
- Labeling of edges:
  Distances \( \Delta x = x_{n'} - x_n \) between nodes \( n \) and \( n' \)
- Graph is not complete

Bunch graph

Constructing a Bunch graph \( B \) from \( M \) Image graphs \( G^M \):
- Summarize the jets from a node \( \rightarrow \) Set of jets \( \rightarrow \) “Bunch”
- Label nodes with Bunches
- Label edges with average distance
  \[ \Delta x_b = \frac{1}{M} \sum_{m \in M} \Delta x^{m}_{mn} \]

Matching

- Goal: Calculate an Image graph for an image
- Four stages:
  1. Find approximate position
  2. Refine position and size
  3. Refine size and find aspect ratio
  4. Local distortion
- Initial Graph: Structure of Bunch Graph
Matching

Top row: Image graphs manually marked
Bottom row: Image graphs found by the system

Comparison of results

- Results of face recognition using the FERET db:
  - Different poses: frontal, halfprofile, profile

<table>
<thead>
<tr>
<th>Method</th>
<th>Model</th>
<th>Gallery</th>
<th>Testfolder</th>
<th>Recognition rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized cross-correlation</td>
<td>194</td>
<td>194</td>
<td>62%</td>
<td></td>
</tr>
<tr>
<td>Linear least squares network</td>
<td>100</td>
<td>100</td>
<td>64%</td>
<td></td>
</tr>
<tr>
<td>Trained matching pursuit filters</td>
<td>311</td>
<td>311</td>
<td>96%</td>
<td></td>
</tr>
<tr>
<td>Principal Component Analysis</td>
<td>150</td>
<td>150</td>
<td>99%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>150 pl</td>
<td>99%</td>
<td></td>
</tr>
<tr>
<td>Elastic Bunch Graph Matching</td>
<td>250</td>
<td>250 pl</td>
<td>97%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>181</td>
<td>97%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>250 pl</td>
<td>86%</td>
<td></td>
</tr>
</tbody>
</table>

Conclusion

- Holistic face recognition
  - Assuming faces are aligned
  - Subspace approach
- Active shape/appearance model
  - Separate shape and appearance
  - Landmark based face warping
- Elastic Bunch Graph Matching
  - Modeling topological with a graph
  - Modeling local appearance with Gabor
- Open problems
  - Alignment
  - Occlusion and cluttering
  - Expression, aging, glasses, facial hair