Face Recognition – Part I

“Fundamentals”

Rainer Stiefelhagen <stiefel@ira.uka.de>

November 27, 2008, Karlsruhe, Germany
Outline

- Cognitive Issues

- Technical Approaches
  - Feature-Based
  - Eigenfaces (PCA) and variants
  - Fisherfaces (LDA)

- (Some applications of face recognition & FR systems under development in cv:hci-group)
Cognitive Issues

- Is face recognition a dedicated process?
  - Face recognition vs. Object recognition
  - Faces are more easily remembered by humans than other objects
  - Prosopagnosia (face blindness) patients are unable to recognize faces, but usually have no other agnosia. They can recognize whether a given object is a face or not.
Is face perception the result of holistic or feature analysis? -Thatcher illusion -

- Global descriptions served as a front end for finer, feature-based perception.
- Face recognition involves more configural/holistic processing than other object recognition.
Which image pairs are from the same person?

Answer: none, these are all different people
Ranking of significance of facial features

- Hair,
- Face outline,
- Eyes and mouth have been found to be important.
- Nose plays an insignificant role in frontal face recognition.
- It could be more important than the eyes or mouth in face recognition using profiles.
Caricatures

- Caricature: “a symbol that exaggerates measurements relative to any measure which varies from one person to another“
- Caricatures do not contain as much information as photographs, but they capture the important characteristics of a face

Al Pacino
Attractiveness & Distinctiveness

- More attractive the faces are, the better is their recognition rate
- The least attractive faces come next
- Followed by the midrange faces
- Humans recognize people from their own race easier than people from another race
- Distinctive faces are better retained in memory and are recognized better and faster than typical faces
The role of spatial frequency analysis
  - Low frequency bands play a dominant role
  - High-frequency components are also required for identification

 Facial Expressions
  - Analysis of facial expressions is accomplished in parallel to face recognition: Some “face blindness“ patients can recognize expressions and patients who suffer from poor expression analysis can perform face recognition quite well.
Viewpoint-invariant recognition?

- Some experiments suggest that memory for faces is highly viewpoint-dependent.
- People are used (trained) to recognize faces from „normal“ illumination
  - with “negative” images, they perform poorly, although, there is the same information
- Movement and face recognition
  - Familiar faces are easier to recognize when shown in moving sequences than in still photographs.
Outline

- Cognitive Issues

- Technical Approaches
  - Feature-Based Approaches
  - Appearance Based Approaches
    - Eigenfaces (PCA) and variants
    - Fisherfaces (LDA)
    - Elastic Bunch Graph Matching - EBMG

- Applications of face recognition
  & FR systems under development in ISL
A Brief History (1900-2000)
Object recognition perspective

- It is not general object recognition!

- It is a single-class object recognition task
Main problem

The variations between the images of the same face due to illumination and viewing direction are almost always larger than image variations due to change in face identity.

-- Moses, Adini, Ullman, ECCV‘94
Source of variations

- Pose
- Illumination
- Occlusion
- Time-gap
- Accessories
- Make-ups
- Aging

1+ years apart
Face Recognition Tasks

- Decision surfaces for three different people
Closed Set vs. Open Set Identification

- **Closed-Set Identification:**
  - The system reports which person from the gallery is shown on the test image: Who is he?
  - Performance metric: Correct identification rate

- **Open-Set Identification:**
  - The system first decides whether the person on the test image is a known or unknown person. If he is a known person who he is?
    1. **False accept:** The invalid identity is accepted as one of the individuals in the database.
    2. **False reject:** An individual is rejected even though he/she is present in the database.
    3. **False classify:** An individual in the database is correctly accepted but misclassified as one of the other individuals in the training data
Authentication/Verification

- A person claims to be a particular member. The system decides if the test image and the training image is the same person: Is he who he claims he is?
- Performance metric: *false reject rate (FRR), false accept rate (FAR)*
- **False reject:** system rejects a valid identity;
- **False accept:** system incorrectly accepts an invalid identity.
Receiver Operating Characteristics (ROC) Curve

- FAR
- FRR
- Forensic applications
- Better system performance
- Equal Error Rate
- Civilian applications
- High security applications
Traditional Approaches

- Feature-based
  - fiducial points
  - distances, angles, areas, etc.
  - Geometrical

- Appearance-based
  - holistic, fiducial regions
  - statistical
Features: Frontal & Profile
Feature-based Face Recognition

- Eyebrow thickness and vertical position at the eye center position
- A coarse description of the left eyebrow‘s arches
- Nose vertical position and width
- Mouth vertical position, width, height upper and lower lips
- Eleven radii describing the chin shape
- Face width at nose position
- Face width halfway between nose tip and eyes

Classification

- Nearest neighbor classifier with Mahalanobis distance as the distance metric.
  \[ \Delta_j(x) = (x - m_j)^T \Sigma^{-1} (x - m_j) \]

  - x: input face image
  - \( m_j \): average vector representing the \( j \)th person.
  - \( \Sigma \): Covariance matrix
  - T: transpose operator

- Different people are characterized only by their average feature vector.

- The distribution is common and estimated by using all the examples in the training set.
Appearance Based Approaches

- can be either
  - holistic, i.e. they process the whole face as the input
  - local / fiducial, i.e. they process facial features, such as eyes, mouth, etc. separately

Brunelli & Poggio (1993)

Eigenfaces

- A face image defines a point in the high dimensional image space
- Different face images share a number of similarities with each other

⇒ They can be described by a relatively low dimensional subspace
⇒ Project the face images into an appropriately chosen subspace and perform classification by similarity computation (distance, angle)
Eigenfaces

- Dimensionality reduction procedure used here is called *Karhunen-Loéve transformation* or *principal component analysis*

- Objective: Find the vectors that best account for the distribution of face images within the entire image space
Principal Component Analysis (PCA)

$y$: face image

$Y = [y_1, y_2, y_3, \ldots, y_K]$  
$Y$: face matrix

$m = (1/K) \times \Sigma y$  
$m$: Mean face

$C = (Y - m)(Y - m)^T$  
$C$: Covariance matrix

$D = U^T C U$  
$D$: eigenvalues, $U$: eigenvectors

$\Omega = U^T \times (y - m)$  
$\Omega$: representation coefficients
Eigenfaces

- **Training:**
  - Acquire initial set of face images (training set):
    - \( Y = [y_1, y_2, y_3, ..., y_K] \)
  - Calculate the eigenfaces from the training set, keeping only the M images corresponding to the highest eigenvalues
    - \( U=(u_1, u_2, .. u_M) \)
  - Calculate representation of each known individual \( k \) in face space
    - \( \Omega_k = U^T * (y_k - m) \)

- **Testing:**
  - Project input new image \( y \) into face space: \( \Omega = U^T * (y - m) \)
  - Find most likely candidate class \( k \) by distance computation
    - \( e_k = || \Omega - \Omega_k || \), for all \( \Omega_k \)
Principal components are the eigenvectors of the covariance matrix of the set of face images

(PCA: see e.g. Duda & Hart, 1973: Pattern Classification, Scene Analysis)
Principal Component Analysis (PCA)
Eigenfaces

Principal components are called “eigenfaces” and they span the “face space”
Projections onto the face space

- Images can be reconstructed by their projections in face space
  - \( Y_f = \sum_{i=1}^{M} \omega_i u_i \)

- Appearance of faces in face-space does not change a lot

- Difference of mean-adjusted image (\( Y - m \)) and projection \( Y_f \) gives a measure of „faceness“
  - \( \rightarrow \) distance from face space (dffs)
  - Can be used to detect faces

Images (left) and their projections in face space (right)
Projections into face space

- Case 1: projection of a *known* individual
  - Near face space ($\varepsilon<\theta_\delta$) and near known face $\Omega_k (\varepsilon_k<\theta_\varepsilon)$

- Case 2: projection of an *unknown* individual
  - Near face space, far from reference vectors

- Case 3 and 4: not a face
  - Far from face space

![Diagram of face space with cases 1 to 4 marked]
Principal Component Analysis (PCA)

Projects all faces onto a universal eigenspace to “encode” via principal components

Uses inverse-distance as a similarity measure \( S(p,g) \) for matching & recognition

Modular Eigenspaces

Does classification using fiducial regions (eyes, nose, mouth is excluded in this study-) instead of using entire face.

View-based Eigenspaces

- Extension: View-based eigenspaces for general viewing conditions

- Given: N individuals under M different views
  - Build universal eigenspace from the combination of NM images: “parametric eigenspace”
  - Build “view-based” set of M separate eigenspaces

- Experiments show slight advantage of view-based approach
View-based Eigenspaces

- Build an eigenspace for each view
- Decide input images’ direction of view using distance from view space metric
- Do classification in that view-space

Bayesian Face Recognition

- Problem: Simple nearest-neighbour similarity measures do not exploit knowledge of critical appearance variations

Bayesian similarity measure:
- denotes belief that image differences are caused by typical appearance variations of an individual (caused by expression, etc.)
- compares typical within-class (intrapersonal) variations with between-class (extrapersonal) variations

Dual PCA (Bayesian)

Intrapersonal $\Omega_I$
Extrapersonal $\Omega_E$

$\Omega_I \equiv \{\Delta = x_i - x_j : L(x_i) = L(x_j)\}$
$\Omega_E \equiv \{\Delta = x_i - x_j : L(x_i) \neq L(x_j)\}$

$$S = P(\Omega_I | \Delta) = \frac{P(\Delta | \Omega_I)P(\Omega_I)}{P(\Delta | \Omega_I)P(\Omega_I) + P(\Delta | \Omega_E)P(\Omega_E)}$$

$P(\Delta | \Omega)$ derived from training data
(some tricks are needed …)

$S(p,g)$ defines a probabilistic similarity metric
Face ID: Eigenfaces

- Problems and shortcomings:
  - Eigenfaces do not distinguish between shape and appearance:
    Active Shape Models (ASM)
    Active Appearance Models (AAM)
  - PCA does not use class information:
    PCA projections are optimal for reconstruction from a
    low dimensional basis, they may not be optimal from
    a discrimination standpoint:
    “Much of the variation from one image to the next is due to
    illumination changes.” [Moses, Adini, Ullman]
Linear Discriminant Analysis (LDA) – Fisherfaces-

- Fischer‘s Linear Discriminant
  - Preserves separability of classes
  - Maximizes ratio of projected between-classes to projected within-class scatter

- Between-class scatter
  \[ S_B = \sum_{i=1}^{c} |x_i|(\mu_i - \mu)(\mu_i - \mu)^T \]

- Within-class scatter
  \[ S_W = \sum_{i=1}^{c} \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T \]

- Where
  - \( c \) is the number of classes
  - \( \mu_i \) is the mean of class \( X_i \)
  - \( |X_i| \) is number of samples of \( X_i \)

\[
W_{fld} = \arg \max_W \frac{W^T S_B W}{W^T S_W W}
\]
FisherFace

\[ S_W = S_1 + S_2 \]

Good separation
Fisherface

Poor Projection

Good Projection
PCA vs. LDA
Fisherfaces

- Fisher’s Linear Discriminant projects away the within-class variation (lighting, expressions) found in training set.

- Fisher’s Linear Discriminant preserves the separability of the classes.

Fisherfaces: Experiments

Subsets of Harvard Database:

Subset 1: angle of light source is within 15 degr. of camera axis

Subset 2: angle of light source is within 30 degr. of camera axis

Subset 3: angle of light source is within 45 degr. of camera axis
Face ID: Fisherfaces

<table>
<thead>
<tr>
<th>Method</th>
<th>Subset 1</th>
<th>Subset 2</th>
<th>Subset 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenface</td>
<td>0.0</td>
<td>4.4</td>
<td>41.5</td>
</tr>
<tr>
<td>Eigenface w/o 1st 3</td>
<td>0.0</td>
<td>4.4</td>
<td>27.7</td>
</tr>
<tr>
<td>Fisherface</td>
<td>0.0</td>
<td>0.0</td>
<td>4.6</td>
</tr>
</tbody>
</table>

Methods were trained on images from Subset 1

(Belhumeur et al., “Eigenfaces vs. Fisherfaces”, TPAMI, 1997)
Applications of Face Recognition

<table>
<thead>
<tr>
<th>Areas</th>
<th>Specific applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entertainment</td>
<td>Video game, virtual reality, training programs</td>
</tr>
<tr>
<td></td>
<td>Human-robot-interaction, human-computer-interaction</td>
</tr>
<tr>
<td>Smart Cards</td>
<td>Drivers’ licenses, entitlement programs</td>
</tr>
<tr>
<td></td>
<td>Immigration, national ID, passports, voter registration</td>
</tr>
<tr>
<td></td>
<td>Welfare fraud</td>
</tr>
<tr>
<td>Information security</td>
<td>TV Parental control, personal device logon, desktop logon</td>
</tr>
<tr>
<td></td>
<td>Application security, database security, file encryption</td>
</tr>
<tr>
<td></td>
<td>Intranet security, internet access, medical records</td>
</tr>
<tr>
<td></td>
<td>Secure trading terminals</td>
</tr>
<tr>
<td>Law enforcement and</td>
<td>Advanced video surveillance, CCTV control</td>
</tr>
<tr>
<td>surveillance</td>
<td>Portal control, postevent analysis</td>
</tr>
<tr>
<td></td>
<td>Shoplifting, suspect tracking and investigation</td>
</tr>
</tbody>
</table>

- Face recognition market totaled revenues worth **$21.5 million in 2002** and is poised to expand to **$791.8 million by 2009** (Frost & Sullivan)
- In ICPR 2004, 44 of ~900 papers were on **face recognition**
FR systems under development in ISL

- Recognizing presenter’s face in smart rooms (CHIL)
- Recognizing lab members (SFB 588- Humanoid Robot)
- Recognizing the person entering the room
- Face recognition for video-retrieval
Recognizing presenter‘s face in the pool-room (CHIL)

Unconstrained face recognition

Adverse illumination conditions due to projector‘s beam

4 camera views
Face Recognition in Smart Room

Cam: cam1
Face resolution: 28x33
Ground truth: AIT_000

AIT_000 29.2674
Recognizing the person entering the room

ICCV’07
Recognizing lab members
(SFB 588- Humanoid Robot)
Human-Robot Interaction
Relevant Papers


Further Reading