Applications

- Law enforcement (mug shot identification)
- Verification for personal identification (driver’s licenses, passports, etc.)
- Surveillance of crowd behavior
- Security applications
- US Visit program

Face is a passive biometric. Does not require cooperation.
Data Collection

Environment: well controlled
• frontal + profile photographs
• uniform background
• identical poses
• similar illumination

Mug-shot

Canonical faces:
cropped, size and position normalized,
minimum background

Face recognition in uncontrolled environment:
• more than 1 face can appear
• lighting conditions vary
• facial expressions
• different scale
• position, orientation
• facial hair, make-up, glasses
• occlusion

Face recognition is a complex problem.

• Detect face
• If multiple, estimate location and size
Approaches

Face Recognition: Representation and Classification

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Variations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensing modality</td>
<td>2-D intensity image, color image, infrared image, 3-D range image, combination of them</td>
</tr>
<tr>
<td>Viewing angle</td>
<td>Frontal views, profile views, general views, or a combination of them</td>
</tr>
<tr>
<td>Temporal component</td>
<td>Static images, time-varying image sequence (may facilitate face tracking, expression identification, etc.)</td>
</tr>
<tr>
<td>Computational tools</td>
<td>programmed knowledge rules, statistical decision rules, neural networks, genetic algorithms, etc.</td>
</tr>
</tbody>
</table>

Imaging Modalities

- Optical Camera (color, black/white)
- Infrared Camera
- Laser radar (new technology)

Image, infrared image, and video sequence
Face Biometric

- **Macro elements**: the mouth, nose, eyes, cheekbones, chin, lips, forehead, and ears.
- **Micro elements**: distances between the macro features or reference features and the size of features.
- **Heat radiation**

**Features derived from face images**: geometric and statistical.

Human can easily detect and identify individual’s face in a scene.

Designing an automated system is hard.

**Block-Diagram**

- **Input Image** ➔ **Face Detection** ➔ **Cropping Normalization** ➔ **Feature Extraction** ➔ **Face Recognition**

Cropping and Normalization Tasks:

1. Facial region extraction (minimizes influence of other factors not related to face);
2. Spatial normalization, which aligns the centers of the eyes and fixes the number of pixels between the eyes via rotation and scaling information;
3. Intensity normalization (converts image into a vector and normalizes it to be zero mean variance one vector.)
Performance Evaluation Factor

- Measured using standard databases and objective performance statistics.
- The face recognition vendor test (FRVT)
- The FRVT 02 reported:
  1. under normal indoor illumination, the state-of-the-art face recognition system achieve 90% verification rate at a false accept rate 1%;
  2. under outdoor conditions, the best vendor can get 50% verification rate at a false accept rate 1%;
  3. the 3-D morphable models techniques improve non-frontal face recognition.
- Illumination and pose are still challenging areas.
- The FRGC06: indoor illumination, 95% verification rate at FAR=0.1%

Face Detection

- Earlier methods: correlation or template matching techniques, matched filters, subspace methods
- Recent methods are data-driven learning-based techniques
  - Statistical modeling (estimation of face – non-face patterns, then apply pattern classifier)
  - Neural network-based learning (learn to discriminate face – non-face patterns using training samples and the network structure)
  - Support Vector Machines
  - Markov Random Field
  - Color based detection
Face Recognition

Requires:
- Low-dimensional representation to achieve data compression
  (Usually starts from dimensionality reduction; high-dimensional space is almost empty).
- Enhanced discrimination abilities
  (Achieving high separability between patterns).

Approaches

Manually defined features
- Geometric features such as distance and angles between geometric points: (ex. eye corners, mouth extremities, nostrils, chin top, etc.)
- For profiles: a set of characteristic points.
- Locations of points can be extracted automatically.

Problems:
- Automatic extraction is not reliable
- The number of features is small
- The reliability of each feature is difficult to estimate
Approaches

Automatically derived features

Nonstatistical Methods:
  Neural networks

Statistical Methods:
  Eigenfaces, nonlinear deformations

Representation and Recognition

<table>
<thead>
<tr>
<th>Techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Representation methods</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Recognition methods</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

October 30, 2007

WVU
Local Feature Analysis
Based on macro features

1. Separation of face from background
2. Reference points are detected used the change in shading around features.
3. Anchor points are tied in triangles.
4. Angles are measured from each of anchor points.
5. 672-bit template is generated.
6. Change in lighting conditions or orientation leads to new templates.
7. Live scan undergoes the same processing. High percentage score results in match.

Eigenfaces
Appearance-based approach

**Eigenface** = “one’s own face”

- Input: 2-D gray scale image
- Image is a high-dimensional vector (each pixel is a component).
- Each image is decomposed in terms of other basis vectors (eigenvectors).

$$ \mathbf{f} = \sum_{k=1}^{N} w_k \mathbf{e}_k $$

Where $N$ is the image dimension, $\mathbf{e}_k$ is the k-th eigenface.
- Template consists of weights $w_k$.

- The features of input image and database templates are compared using nearest neighbor rule (ex. 1-NN = Euclidean distance).
Neural Network: Detection

- **Training Set:** N face images with identified macro features are fed into network + other random images.
- Other faces are entered with no identified macro features.
- The unidentified faces are re-entered into system with identified features.

The parts of ANN: (a) face detection and framing; (b) ANN input level; (c) Receptive fields; (d) Hidden units; (e) Output.

Neural Network

**Face Detection and Framing:**
face is separated from its background, framed, and transformed into appropriate size.

**ANN input level:**
face image is converted into pixels to correspond to array of input neurons.

**Receptive fields:**
the mapping is chosen to reflect general characteristics of face

**Hidden units:**
have a one-to-one neuron/receptive field relationship. Hidden units determine if appropriate feature was detected.

**Output:**
a single output neuron that indicates positive or negative face match.
Representation and Recognition

- **Principle Component Analysis (PCA)** derives an orthogonal projection basis that leads directly to dimensionality reduction and feature selection.
  - Eigenfaces = eigenvectors related to the largest eigenvalues
  - PCA is optimal criterion for dimensionality reduction, but **does not always provide good discrimination.**
    [Kirby90, Turk91]
- Solution: Integrate PCA with the Bayes classifier (Probabilistic Reasoning Models) [Liu00]
- **Shape and texture** applies a two-stage procedure
  - Coding starts by marking important internal and boundary points.
  - The points are aligned using translation, scaling, and rotation.
  - Calculate the average of control points – defines the mean shape.
  - Triangulate the marked face and warp each face into the mean shape.
  - The first stage yields the shape, the second stage is related to texture. [Beymer95, Cootes98, Craw92, Edwards98, Lanitis97, Liu01]

Representation and Recognition

- **Classifiers used:**
  - Mahalanobis distance classifier [Lanitis97] from shape and texture.
  - Enhanced Fisher classifier [Liu01]
- **The Gabor wavelets:** have desirable characteristics of spatial locality and orientation selectivity.
  - Computation of Gabor features and use of flexible graph matching [Lades93]
  - Gabor filters were used for two-class categorization of gender, race, and facial expression [Lyons00]
  - Steps include registration of grid with the face using elastic matching and characterization of extracted points using LDA.
  - Gabor-Fisher classifier (robust to illumination and facial expression variability) [Liu02]
Representation and Recognition

- The Bayes classifier is optimal when the probability distributions of data are known.
- The PRM method [Liu00] integrates PCA and Bayes classifier.
- Fisher Linear Discriminant (FLD, LDA): separates the mean of different classes as far as possible and compresses the same class data as much as possible.
- Independent Component Analysis (ICA), a technique for blind source separation.
- ICA analysis, when carried out in the properly compressed and whitened space, performs better than the Eigenface and Fisherface methods.
- When augmented by additional criteria (such as the MAP rule) its performance degrades.
- Graph matching achieves invariance to affine transformations or localized facial expressions. Graph nodes have to be manually defined to find the corresponding nodes in the different graphs.

Kernel-Based Methods

- Kernel PCA, kernel FLD, and SVM overcome the limitations of the linear approaches by nonlinearly mapping the input space into a high-dimensional feature space.
- T. Cover: Non-linearly separable patterns in an input space are linearly separable with high probability if the input space is transformed nonlinearly to a high-dimensional feature space.
- 3D methods provide potential solutions to pose invariant recognition. 3D models are often derived from laser-scanned 3D heads (range data) or reconstructed using shape from shading.
- Hsu and Jain [Hsu01] described the method of modeling 3D faces based on a triangular mesh model and individual facial measurements. A potential solution to face recognition with variations in illumination, pose, and facial expressions.
- Method by Zhao and Chellappa [Zhao01] applies a 3D model to synthesize a prototype image from a given image acquired under different lighting and viewing conditions.
Face: Pros and Cons

Pros:

• Used for manual inspection: driver license, passport. Wide public acceptance for this biometric identifier.
• The least intrusive from sampling point of view, requiring no contact.
• Face recognition can be used (at least in theory) for screening of unwanted individuals in a crowd, in real time.
• It is a good biometric identifier for small-scale verification applications.

Cons:

• For robust identification, face needs to be well lighted by controlled source.
• Currently it performs poor in outdoor protocol.
• Disguise is an obvious circumvention method. Disguised person is not identified.
• There is some criminal association with face identifiers since it has been used by law enforcement agencies (“mug-shots”).
• Privacy concerns.

Face Databases

• The Olivetti (ORL, now AT&T) database: (40 subjects, ten 92x112 pixels with a variety of lighting and facial expressions)
  http://www.uk.research.att.com/facialrecognition
• FERET (14,126 images that includes 1,199 subjects and 356 duplicate sets)
  http://www.dodcounterdrug.com/facialrecognition
• FRVT 2002 (120,000 faces, includes video of faces)
  http://www.frvt.org/FRVT2002/default.htm
• NIST 18 Mugshot Identification Database (3,248 mugshot images: front images and profiles, 500 dpi)
• The MIT database (16 subjects, 27 images per subject with varying illumination, scale, and head orientation)
  ftp://whitechapel.media.mit.edu/pub/images/
• The Yale database (5,850 images of 10 subjects each imaged under 576 viewing conditions: 9 poses and 64 illumination conditions. Size 640x480, 256 grey levels.)
  http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html
• The Purdue database (4,000 color images from 126 subjects imaged with different expressions, illumination conditions, and occlusion.)
  http://rvl1.ecn.purdue.edu/aleix/aleix_face_DB.html
References