Announcements

- HW3 assigned, due on Thursday.
- Last HW fingerprint recognition – to be assigned next week.
- Thursday – guest lecture
  “Forensic DNA Technology”
  Dr. David King, IntegenX, Executive Vice President, Product Development
  – Attendance is mandatory

HW3 Hand Recognition Challenge

- Data: Your hand outlines
- Goal: Build two different classifiers
- Features: Manual
- Evaluation:
  1. Four-fold cross validation on training set
  2. Unlabeled test set
- Winner – 10 pts extra credit on HW.

Hand outline acquisition

- The Hand Recognition Challenge
  You are the data
- You will be given 5 pieces of paper.
- You will be assigned an ID number
- On the upper right corner, write
  <Your ID> - 1 ….. <Your ID> - 5
- Flip over pages, so ID is on back
- Trace your left hand on four pieces of paper.
- Now have your neighbor trace your left hand on the fifth pieces of paper.

Image as a Feature Vector

- Consider an n-pixel image to be a point in an n-dimensional space, \( x \in \mathbb{R}^n \).
- Each pixel value is a coordinate of \( x \).

Eigenfaces: linear projection

- An n-pixel image \( x \in \mathbb{R}^l \) can be projected to a low-dimensional feature space \( y \in \mathbb{R}^k \) by
  \[ y = Wx \]
  where \( W \) is an \( n \times m \) matrix.
- Recognition is performed using nearest neighbor in \( \mathbb{R}^k \).
- How do we choose a good \( W \)?
Eigenfaces: Principal Component Analysis (PCA)

Assume we have a set of n feature vectors $x_i$ (i = 1, ..., n) in $\mathbb{R}^d$. Write

$$\mu = \frac{1}{n} \sum x_i,$$

$$\Sigma = \frac{1}{n-1} \sum (x_i - \mu)(x_i - \mu)^T$$

The unit eigenvectors of $\Sigma$—which we write as $v_1, v_2, \ldots, v_n$—where the order is given by the size of the eigenvalue and $v_1$ has the largest eigenvalue—give a set of features with the following properties:

- They are independent.
- Projection onto the basis $\{v_1, \ldots, v_n\}$ gives the $d$-dimensional set of linear features that preserves the most variance.

Some details: Use Singular value decomposition, "trick" described in text to compute basis when n<<d

How do you construct Eigenspace?

$$[x_1, x_2, x_3, x_4, x_5]$$

Construct data matrix by stacking vectorized images and then apply Singular Value Decomposition (SVD)

Matrix Decompositions

- Definition: The factorization of a matrix $M$ into two or more matrices $M_1, M_2, \ldots, M_p$ such that $M = M_1M_2\ldots M_p$.
- Many decompositions exist…
  - QR Decomposition
  - LU Decomposition
  - LDU Decomposition
  - Etc.

SVD Properties

- In Matlab $[u \ s \ v] = svd(A)$, and you can verify that: $A=UsV^T$
- $r=\text{Rank}(A) = \#$ of non-zero singular values.
- $U, V$ give an orthonormal bases for the subspaces of $A$:
  - 1st $r$ columns of $U$: Column space of $A$
  - Last $m-r$ columns of $U$: Left nullspace of $A$
  - 1st $r$ columns of $V$: Row space of $A$
  - Last $n-r$ columns of $V$: Nullspace of $A$
- For some $d$ where $d \leq r$, the first $d$ column of $U$ provide the best $d$-dimensional basis for columns of $A$ in least squares sense.

Singular Value Decomposition

- Any $m \times n$ matrix $A$ may be factored such that $A = USV^T$ where $U$ has $m$-by-$n$ matrix and $V$ has $n$-by-$n$ matrix, and $S$ is a diagonal matrix with non-negative entries $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_s$ with $s = \min(m,n)$ are called the singular values.
- Columns of $U$ are the eigenvectors of $AA^T$ and columns of $V$ are the eigenvectors of $A^TA$!!
- Important property:
  - Singular values are the square roots of Eigenvalues of both $AA^T$ and $A^TA$ & Columns of $U$ are corresponding Eigenvectors!!

Performing PCA with SVD

- Singular values of $A$ are the square roots of eigenvalues of both $AA^T$ and $A^TA$ & Columns of $U$ are corresponding Eigenvectors
- And $\sum \sigma_i a_i = [a_1, a_2, \ldots, a_n]^T [a_1, a_2, \ldots, a_n]^T = AA^T$
- Covariance matrix is:
  $$\Sigma = \frac{1}{n} \sum (x_i - \mu)(x_i - \mu)^T$$
- So ignoring $1/n$, subtract mean image $\mu$ from each input image, create a $d \times n$ data matrix, and perform thin SVD on the data matrix. $D=[x_1^T \mu | x_2^T \mu | \ldots | x_n^T \mu]$
Thin SVD

- Any \( m \times n \) matrix \( A \) may be factored such that
  \[
  A = U \Sigma V^T
  \]
  \[\begin{bmatrix}
  m \\
  x \\
  n 
  \end{bmatrix}
  = \begin{bmatrix}
  m \\
  x \\
  m \\
  \end{bmatrix} \begin{bmatrix}
  m \\
  x \\
  n 
  \end{bmatrix} \begin{bmatrix}
  n \\
  x \\
  n 
  \end{bmatrix}
  \]

- If \( m > n \), then one can view \( \Sigma \) as: (i.e., more pixels than images)
  \[
  \Sigma = \begin{bmatrix}
  \Sigma \\
  0
  \end{bmatrix}
  \]

  Where \( \Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_s) \) with \( s = \min(m,n) \), and lower matrix is \( (n-m \times m) \) of zeros.

- Alternatively, you can write:
  \[
  A = U' \Sigma V'
  \]

- In Matlab, thin SVD is:
  \[
  [U S V] = \text{svd}(A,'econ')
  \]

Eigenfaces

- Modeling
  1. Given a collection of \( n \) labeled training images,
  2. Compute mean image and covariance matrix.
  3. Compute \( k \) Eigenvectors (note that these are images) of covariance matrix corresponding to \( k \) largest Eigenvalues.
  4. Project the training images to the \( k \)-dimensional Eigenspace.

- Recognition
  1. Given a test image, project to Eigenspace.
  2. Perform classification to the projected training images.

Eigenfaces: Training Images

Variable Lighting
Reconstruction using Eigenfaces

- Given image on left, project to Eigenspace, then reconstruct an image (right).

Underlying assumptions

- Background is not cluttered (or else only looking at interior of object)
- Lighting in test image is similar to that in training image.
- No occlusion
- Size of training image (window) same as window in test image.

Difficulties with PCA

- Projection may suppress important detail
  - smallest variance directions may not be unimportant
- Does not generalize well to unseen conditions
- Method does not take discriminative task into account
  - typically, we wish to compute features that allow good discrimination
  - not the same as largest variance

Illumination Variability

"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

Fisherfaces: Class Specific Linear Projection


- An n-pixel image $x \in \mathbb{R}^n$ can be projected to a low-dimensional feature space $y \in \mathbb{R}^m$ by $y = Wx$

where $W$ is an n by m matrix.

- Recognition is performed using nearest neighbor in $\mathbb{R}^m$.
- How do we choose a good $W$?
Covariance of projected samples

Let
- \( \Sigma_x \) be a covariance matrix
- \( W \) be an \( n \times m \) matrix defining a linear projection \( y = Wx \)
- The mean of \( y \) is given by \( \mu_y = W\mu_x \)

Then the covariance of \( y \) is given by:

\[
\Sigma_y = W^T \Sigma_x W
\]

PCA & Fisher’s Linear Discriminant

• Between-class scatter
  \[ S_B = \sum \sum (\mu_i - \mu)(\mu_i - \mu)^T \]
• Within-class scatter
  \[ S_W = \sum \sum (\mu_k - \mu)^T(\mu_k - \mu) \]
• Total scatter
  \[ S_T = \sum \sum (\mu_k - \mu)(\mu_k - \mu)^T = S_B + S_W \]
• Where
  - \( c \) is the number of classes
  - \( \mu_i \) is the mean of class \( \chi_i \)
  - \( |\chi_i| \) is number of samples of \( \chi_i \).

Computing the Fisher Projection Matrix

\[
W_F = \arg \max_W \frac{\|W^T S_B W\|}{\|W^T S_W W\|}
\]

where \( \{w_i \} \) is the set of generalized eigenvectors of \( S_B \) and \( S_W \) corresponding to the \( m \) largest generalized eigenvalues \( \lambda_i \), i.e.,

\[
S_W w_i = \lambda_i S_B w_i, \quad i = 1, 2, \ldots, m
\]

• The \( w_i \)'s are orthonormal
• There are at most \( c-1 \) non-zero generalized Eigenvalues, so \( m \leq c-1 \)
• Can be computed with \texttt{eig} in Matlab

Fisherfaces

\[
W = W_F W_{PCA}
\]

\[
W_{PCA} = \arg \max_W \|W^T S_B W\|
\]

\[
W_F = \arg \max_W \frac{\|W^T S_B W\|}{\|W^T S_W W\|}
\]

• Since \( S_W \) is rank \( N-c \), project training set to subspace spanned by first \( N-c \) principal components of the training set.
• Apply FLD to \( N-c \) dimensional subspace yielding \( c-1 \) dimensional feature space.

• 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.
**Experimental Results - 1**

Variation in Facial Expression, Eyewear, and Lighting

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

**Input:**
- With glasses
- Without glasses

**3 Lighting conditions**
- 5 expressions

**FLD For Glasses/No Glasses Recognition**

**Effect of Cropping**

**Experimental Results - 2**

**Harvard Face Database**

- 10 individuals
- 66 images per person
- Train on 6 images at 15°
- Test on remaining images
**Recognition Results: Lighting Extrapolation**

- **FERET 2000 database**
  - Best ID rate: eigenfaces 80.0%, fisherface 93.2%

**FACE RECOGNITION: STATE OF THE ART?**

From Executive Summary

"Using the most accurate face recognition algorithm, the chance of identifying the unknown subject (at rank 1) in a database of 1.6 million criminal records is about 92%.

"Facial recognition algorithms are more accurate on the visa images than the mug shot images.

"The visa images are collected with careful cooperation of the subject, active compliance by the photographer to the image collection specification, and a yes/no review by an official."

**UNCONSTRAINED FACE RECOGNITION**

- Pose
- Illumination
- Expression
- Aging
- Distress
- Glasses & sunglasses
- Facial hair
- Headgear
- Occlusion
- Blur
- Resolution
- Low quality

**A PROGRESSION OF BENCHMARKS**

- **FERET** -- Well controlled images
- **Yale Face Database B** -- Systematic variation of pose & lighting
- **Multi-PIE** -- Systematic variation in pose, lighting, expression, multiple sessions
- **NIST Multiple Biometric Evaluation** -- Large scale, mostly well-controlled
- **Labeled Faces in the Wild (LFW)** -- Toward unconstrained
Labeled Faces in the Wild: De Facto Standard for Evaluation

- 13,233 images of 5,749 individuals
- Acquired from Yahoo News in 2002
- Results reported for 38 commercial and academic methods

3000 “same” pairs
3000 “different” pairs

G. Huang, M. Ramesh, T. Berg, E. Learned-Miller, 2007

Toward unconstrained face recognition

Attribute and Simile Classifiers for Face Verification

ICCV 2009
Neeraj Kumar
Alexander C. Berg
Peter N. Belhumeur
Shree K. Nayar
Columbia University

Attributes can define categories

- Female
- Eyeglasses
- Middle-aged
- Dark hair

- Caucasian
- Teeth showing
- Outside
- Tilted head

Are these images of the same person?

http://mughunt.securics.com/
Prior approaches

Images $\rightarrow$ Low-level features $\rightarrow$ Verification

RGB
HOG
LBP
SIFT
...
RGB
HOG
LBP
SIFT
...

Different

Our approach: attributes

Images $\rightarrow$ Low-level features $\rightarrow$ Attributes $\rightarrow$ Verification

RGB
HOG
LBP
SIFT
...
RGB
HOG
LBP
SIFT
...

Male
Asian
Dark hair
Round jaw
Different

PubFig dataset & benchmark

Public figures:
• Politicians
• Celebrities
Larger & deeper:
• 60,000 Images
• 200 People
• 300 Images per person
Subsets:
• Pose
• Illumination
• Expression

http://www.cs.columbia.edu/CAVE/databases/pubfig/

Amazon Mechanical Turk

500,000 Attribute Labels = $5,000 + 1 month
See also [Deng, et al., 2009] [Vijayanarasimhan & Grauman, 2009]

Learning an attribute classifier

Training images $\rightarrow$ Low-level features $\rightarrow$ Feature selection $\rightarrow$ Train classifier

Males

RGB
HoG
HSV
...
RGB, Nose
HoG, Eyes
HSV, Hair
Edges, Mouth
...

Females

Gender classifier
Male
0.87

Using attributes to perform verification

Verification classifier
Previous state-of-the-art on LFW

Atributes for Recognition on LFW

HUMAN FACE VERIFICATION PERFORMANCE AND CONTEXT

LFW Results May 2014