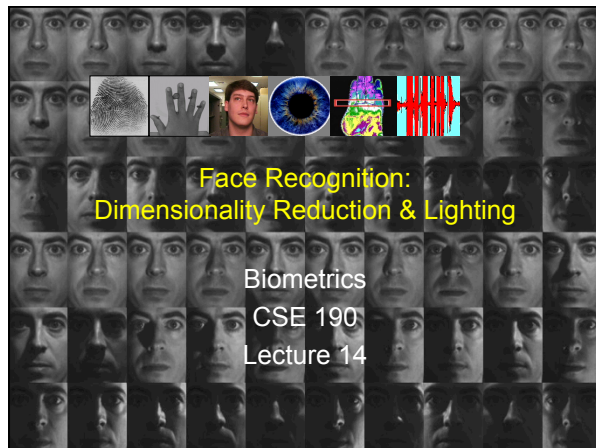




- Literature review due on Today
- Hand classification will be assigned on Tuesday.

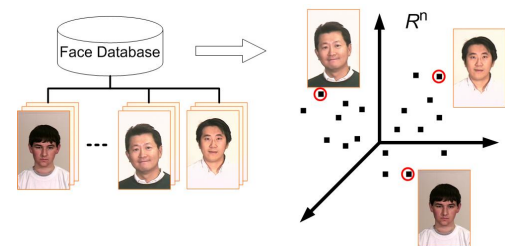
CSE190, Winter 2011

The Space of Images

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

UCSD Jacobs Computer Science and Engineering

- Face Space:
 - A set of face images construct a *face space* in $\mathbb{R}^n$
 - Appearance-based methods analyze the distributions of individual faces in face space



CS252A, Winter 2011

Computer Vision I

Eigenfaces: linear projection

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

$$\mathbf{y} = \mathbf{W}\mathbf{x}$$
 where $\mathbf{W}$ is an n by m matrix.
- Recognition is performed using nearest neighbor in $\mathbb{R}^m$.
- How do we choose a good $\mathbf{W}$?

CS252A, Winter 2011

Computer Vision I

Eigenfaces: Principal Component Analysis (PCA)

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

$$\mu = \frac{1}{n} \sum_i \mathbf{x}_i$$

$$\Sigma = \frac{1}{n-1} \sum_i (\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)^T$$

The unit eigenvectors of Σ — which we write as $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_d$, where the order is given by the size of the eigenvalue and $\mathbf{v}_1$ has the largest eigenvalue — give a set of features with the following properties:

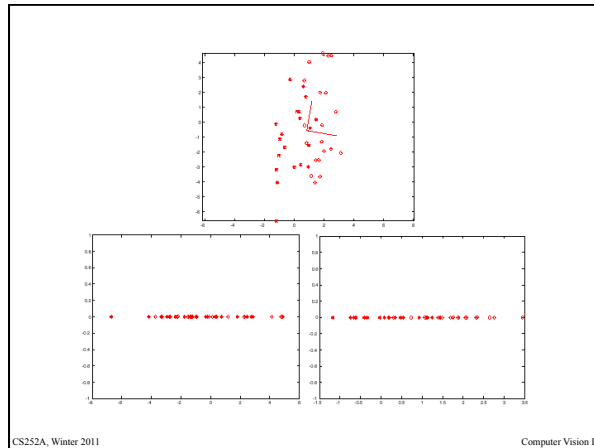
- They are independent.
- Projection onto the basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ gives the k -dimensional set of linear features that preserves the most variance.

Algorithm 22.5: Principal components analysis identifies a collection of linear features that are independent, and capture as much variance as possible from a dataset.

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

CS252A, Winter 2011

Computer Vision I



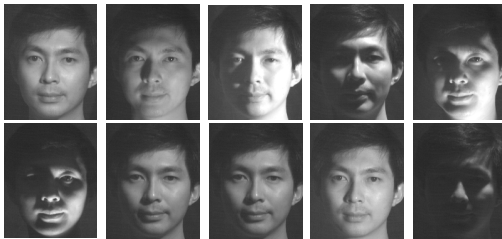
Eigenfaces

- Unsupervised dimensionality reduction
- Subspace that maximizes variance of projected samples and consequently minimizes reconstruction error.
- Computation using eigenstructure of covariance matrix, SVD on data matrix, or thin SVD on data matrix when there are fewer samples than number of pixels.
- Classification within subspace (nearest neighbor, parametric models, non-parametric methods)
- Face detection with distance to subspace.

CS252A, Winter 2011

Computer Vision I

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

CS252A, Winter 2011

Computer Vision I

Fisherfaces: Class Specific Linear Projection

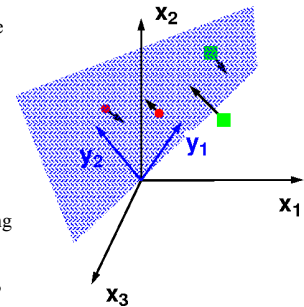
P. Belhumeur, J. Hespanha, D. Kriegman, *Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*, PAMI, July 1997, pp. 711–720.

- 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 ?



CS252A, Winter 2011

Computer Vision I

Covariance of projected samples

Let

- Σ_x be a covariance matrix
- W be an n by m matrix defining a linear projection $y=Wx$

- The mean of y is given by

$$\mu_y = W\mu_x$$

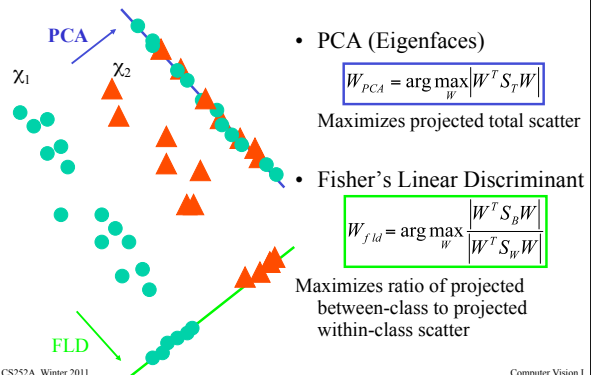
Then the covariance of y is given by:

$$\Sigma_y = W^T \Sigma_x W$$

CS252A, Winter 2011

Computer Vision I

PCA & Fisher's Linear Discriminant



CS252A, Winter 2011

Computer Vision I

PCA & Fisher's Linear Discriminant

- Between-class scatter

$$S_B = \sum_{i=1}^c |\chi_i| (\mu_i - \mu)(\mu_i - \mu)^T$$

- Within-class scatter

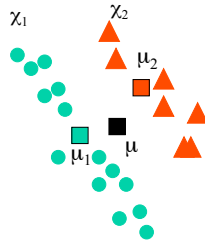
$$S_W = \sum_{i=1}^c \sum_{x_i \in \chi_i} (x_i - \mu_i)(x_i - \mu_i)^T$$

- Total scatter

$$S_T = \sum_{i=1}^c \sum_{x_i \in \chi_i} (x_i - \mu)(x_i - \mu)^T = S_B + S_W$$

- Where

- c is the number of classes
- μ_i is the mean of class χ_i
- $|\chi_i|$ is number of samples of χ_i .



CS252A, Winter 2011

Computer Vision I

PCA & Fisher's Linear Discriminant

PCA

χ_1

χ_2

μ_1

μ_2

μ

FLD

- PCA (Eigenfaces)

$$W_{PCA} = \arg \max_W |W^T S_T W|$$

Maximizes projected total scatter

- Fisher's Linear Discriminant

$$W_{FLD} = \arg \max_W \frac{|W^T S_B W|}{|W^T S_W W|}$$

Maximizes ratio of projected between-class to projected within-class scatter

CS252A, Winter 2011

Computer Vision I

Computing the Fisher Projection Matrix

$$W_{opt} = \arg \max_W \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1 \ w_2 \ \dots \ w_m] \quad (4)$$

where $\{w_i | i = 1, 2, \dots, m\}$ is the set of generalized eigenvectors of S_B and S_W corresponding to the m largest generalized eigenvalues $\{\lambda_i | i = 1, 2, \dots, m\}$, i.e.,

$$S_B w_i = \lambda_i S_W w_i, \quad i = 1, 2, \dots, m$$

- 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 *eig* in Matlab

CS252A, Winter 2011

Computer Vision I

Fisherfaces

$$W = W_{FLD} W_{PCA}$$

$$W_{PCA} = \arg \max_W |W^T S_T W|$$

$$W_{FLD} = \arg \max_W \frac{|W^T W_{PCA}^T S_B W_{PCA} W|}{|W^T W_{PCA}^T S_W W_{PCA} W|}$$

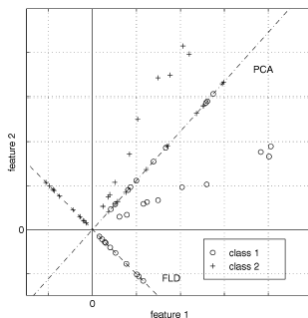
- Since S_W is rank $N-c$, project training set to subspace spanned by first $N-c$ principal components of the training set.
- Apply FLD to $N-c$ dimensional subspace yielding $c-1$ dimensional feature space.

- 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.

CS252A, Winter 2011

Computer Vision I

PCA vs. FLD



CS252A, Winter 2011

Computer Vision I

Experimental Results - 1

Variation in Facial Expression, Eyewear, and Lighting

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

With glasses Without glasses 3 Lighting conditions 5 expressions



CS252A, Winter 2011

Computer Vision I

FLD For Glasses/No Glasses Recognition

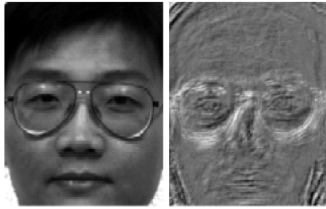


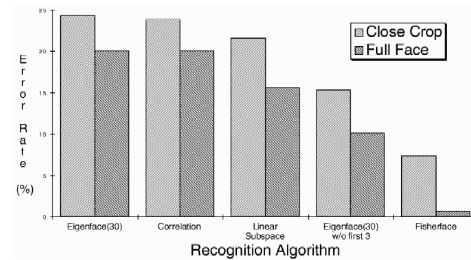
TABLE 1
COMPARATIVE RECOGNITION ERROR RATES FOR GLASSES/
NO GLASSES RECOGNITION USING THE YALE DATABASE

Glasses Recognition		
Method	Reduced Space	Error Rate (%)
PCA	10	52.6
Fisherface	1	5.3

CS252A, Winter 2011

Computer Vision I

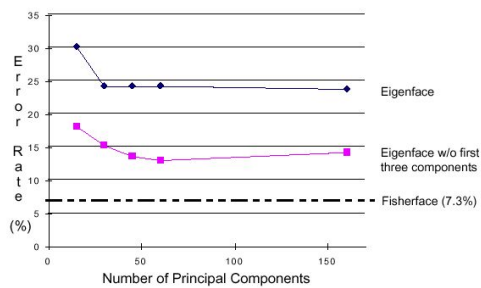
Effect of Cropping



CS252A, Winter 2011

Computer Vision I

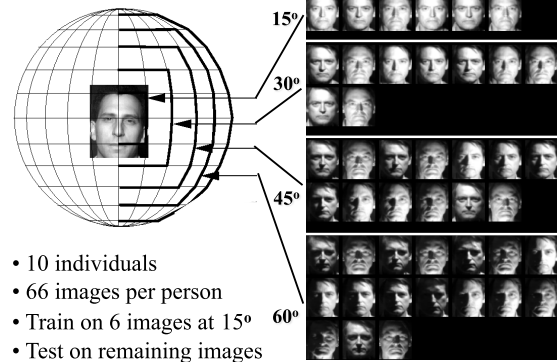
Experimental Results - 2



CS252A, Winter 2011

Computer Vision I

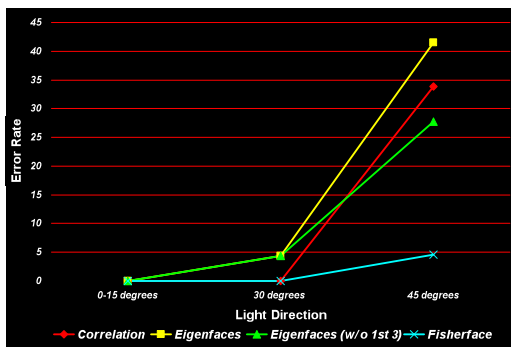
Harvard Face Database



CS252A, Winter 2011

Computer Vision I

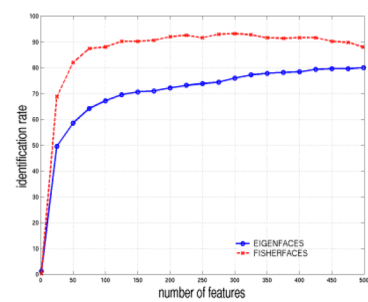
Recognition Results: Lighting Extrapolation



CS252A, Winter 2011

Computer Vision I

• FERET 2000 database



Best ID rate: eigenfaces 80.0%, fisherfaces 93.2%

CS252A, Winter 2011

Computer Vision I

(2D) Model-based – Active Appearance Model

• Model Construction (linear)



labeled image

landmarks

shape-free texture

© Jain, 2004

Active Appearance Model (AAM)

• Shape model

$$s = (x_1, y_1, \dots, x_n, y_n)^T \quad s = \bar{s} + P_s b_s$$

• Appearance model

$$g = (I_1, \dots, I_m)^T \quad g = \bar{g} + P_g b_g$$

• Combined model

$$b = \begin{pmatrix} W_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} W_s P_s^T (s - \bar{s}) \\ P_g^T (g - \bar{g}) \end{pmatrix} \xrightarrow{\text{PCA}} b = Qc$$

© Jain, 2004

AAM

• Model fitting

- Minimize the objective function (gray level difference between the given image and the stored model)

$$\Delta = |\delta I|^2$$

– Searching by learning

- Annotated model (true model parameters)
- Relation: known model displacements $\leftrightarrow$ observed difference vector
- Use multivariate multiple regression to learn the relation and predict the displacement during searching

© Jain, 2004

AAM



Initial
Original

3 its

8 its

11 its

Converged

© Jain, 2004