Eigenimages

- Unitary transforms
- Karhunen-Loève transform
- Eigenimages for recognition
- Sirovich and Kirby method
- Example: eigenfaces
- Eigenfaces vs. Fisherfaces

Unitary transforms

- Sort samples $f[x,y]$ of an MxN image (or a rectangular portion of the image) into column vector of length MN
- Compute transform coefficients

$$\vec{c} = A\vec{f}$$

where $\boldsymbol{A}$ is a matrix of size MNxMN

- The transform A is unitary, iff

$$A^{-1} = \underbrace{A^{*T} \equiv A^{H}}_{\text{Hermitian conjugate}}$$

- If A is real-valued, i.e., $A=A^*$, transform is „orthonormal"

Energy conservation with unitary transforms

- For any unitary transform $\vec{c} = A\vec{f}$ we obtain

$$\|\vec{c}\|^2 = \vec{c}^H \vec{c} = \vec{f}^H A^H A\vec{f} = \|\vec{f}\|^2$$

- Interpretation: every unitary transform is simply a rotation of the coordinate system (and, possibly, sign flips)
- Vector length („energy“) is conserved.

Energy distribution for unitary transforms

- Energy is conserved, but, in general, unevenly distributed among coefficients.
- Autocorrelation matrix

$$R_{cc} = E\left[\vec{c}\vec{c}^{H}\right] = E\left[A\vec{f} \cdot \vec{f}^{H} A^{H}\right] = AR_{ff}A^{H}$$

- Mean squared values („average energies") of the coefficients c_i „live" on the diagonal of R_{cc}

$$E\left[c_i^2\right] = \left[R_{cc}\right]_{i,i} = \left[AR_{ff}A^{H}\right]_{i,i}$$

Eigenmatrix of the autocorrelation matrix

Definition: eigenmatrix $\boldsymbol{\Phi}$ of autocorrelation matrix R_{ff}

- $\boldsymbol{\Phi}$ is unitary
- The columns of $\boldsymbol{\Phi}$ form a set of eigenvectors of R_{ff}, i.e.,

$$R_{ff}\Phi = \Phi\Lambda$$

← Λ is a diagonal matrix of eigenvalues λ_i

$$\Lambda = \begin{pmatrix} \lambda_0 & & & 0 \\ & \lambda_1 & & \\ & & \ddots & \\ 0 & & & \lambda_{MN-1} \end{pmatrix}$$

- R_{ff} is symmetric nonnegative definite, hence $\lambda_i \geq 0 \quad \text{for all } i$
- R_{ff} is normal matrix, i.e., $R_{ff}^H R_{ff} = R_{ff} R_{ff}^H$, hence unitary eigenmatrix exists („spectral theorem“)

Karhunen-Loève transform

- Unitary transform with matrix

$$A = \Phi^H$$

where the columns of Φ are ordered according to decreasing eigenvalues.

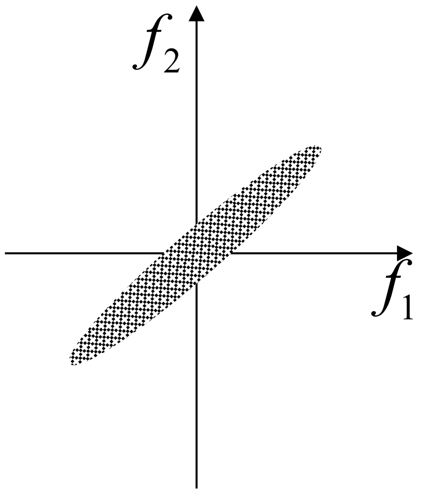
- Transform coefficients are pairwise uncorrelated

$$R_{cc} = AR_{ff}A^H = \Phi^H R_{ff}\Phi = \Phi^H \Phi\Lambda = \Lambda$$

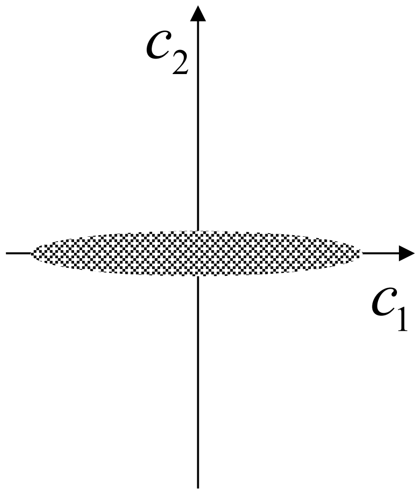
- Energy concentration property:
  - Mean squared approximation error by choosing only first J coefficients is minimized.
  - No other unitary transform packs as much energy into the first J coefficients, for any J

Illustration of energy concentration

Strongly correlated samples,
equal energies



$$\mathbf{A} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$



<u>After KLT:</u>
uncorrelated samples,
most of the energy in
first coefficient

Basis images and eigenimages

- For any unitary transform, the inverse transform

$$\vec{f} = A^H \vec{c}$$

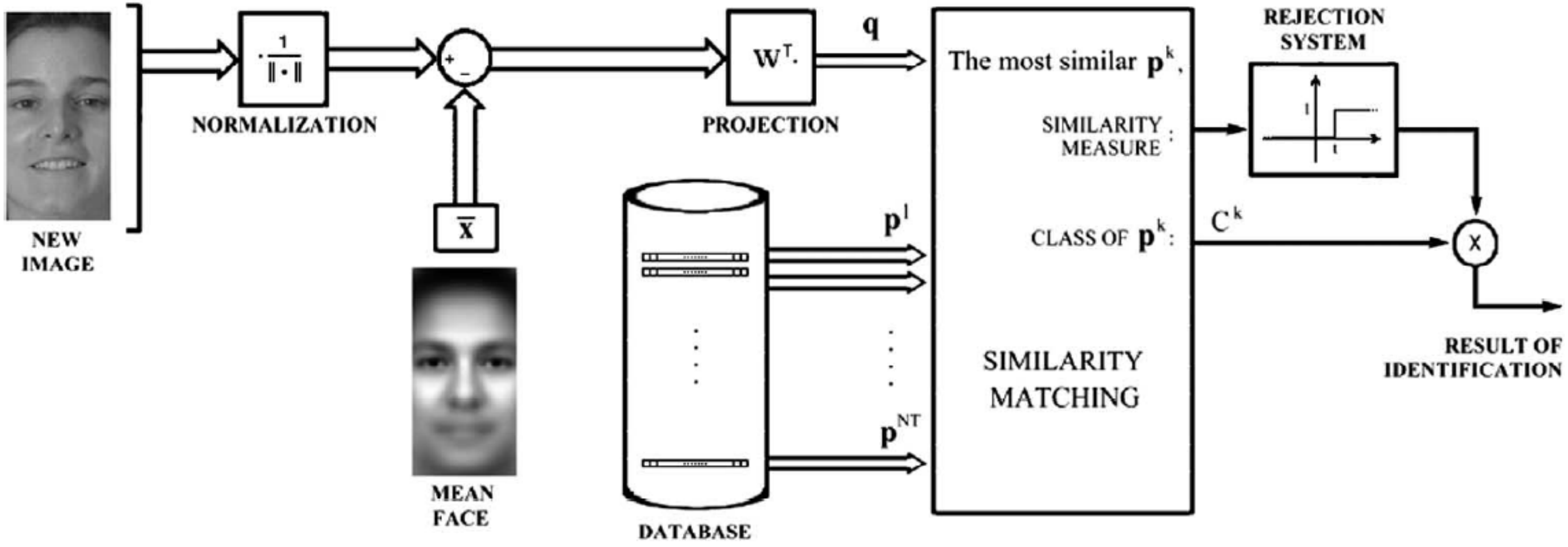
  can be interpreted in terms of the superposition of „basis images“ (columns of A^H) of size MN.
- If the transform is a KL transform, the basis images (aka eigenvectors of the autocorrelation matrix R_{ff}) are called „eigenimages.“
- If energy concentration works well, only a limited number of eigenimages is needed to approximate a set of images with small error. These eigenimages form an optimal linear subspace of dimensionality J.

Eigenimages for recognition

- To recognize complex patterns (e.g., faces), large portions of an image (say of size MN) might have to be considered
- High dimensionality of "image space" means high computational burden for many recognition techniques

  Example: nearest-neigbor search requires pairwise comparison with every image in a database
- Transform $\vec{c} = W\vec{f}$ can reduce dimensionality from MN to J by representing the image by J coefficients
- Idea: tailor a KLT to the specific set of images of the recognition task to preserve the salient features

Eigenimages for recognition (cont.)



[Ruiz-del-Solar and Navarrete, 2005]

Computing eigenimages from a training set

- How to measure $MNxMN$ covariance matrix?
  - Use training set $\vec{\Gamma}_1, \vec{\Gamma}_2, \ldots, \vec{\Gamma}_L$ (each column vector represents one image)
  - Let $\vec{\mu}$ be the mean image of all samples
  - Define training set matrix $S = \left(\vec{\Gamma}_1 - \vec{\mu}, \vec{\Gamma}_2 - \vec{\mu}, \vec{\Gamma}_3 - \vec{\mu}, \ldots, \vec{\Gamma}_L - \vec{\mu} \right)$,

    and calculate $R = \sum_{l=1}^{L} \left(\vec{\Gamma}_l - \vec{\mu} \right) \left(\vec{\Gamma}_l - \vec{\mu} \right)^H = SS^H$

  Problem 1: Training set size should be $L >> MN$

  If $L < MN$, covariance matrix R is rank-deficient

  Problem 2: Finding eigenvectors of an $MNxMN$ matrix.

- Can we find a small set of the most important eigenimages from a small training set $L << MN$?

Sirovich and Kirby method

- Instead of eigenvectors of SS^H, consider the eigenvectors of $S^H S$, i.e.,

  $$S^H S\vec{v}_i = \lambda_i \vec{v}_i$$

  Premultiply both sides by S

  $$SS^H S\vec{v}_i = \lambda_i S\vec{v}_i$$

  By inspection, we find that $S\vec{v}_i$ are eigenvectors of SS^H

- For $L << MN$ this gives rise to great computational savings, by
  - Computing the LxL matrix $S^H S$
  - Computing L eigenvectors $\vec{v}_i$ of $S^H S$
  - Computing eigenimages corresponding to the $L_0 \leq L$ largest eigenvalues according as $S\vec{v}_i$

L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," Journal of the Optical Society of America A, 4(3), pp. 519-524, 1987.

Example: eigenfaces

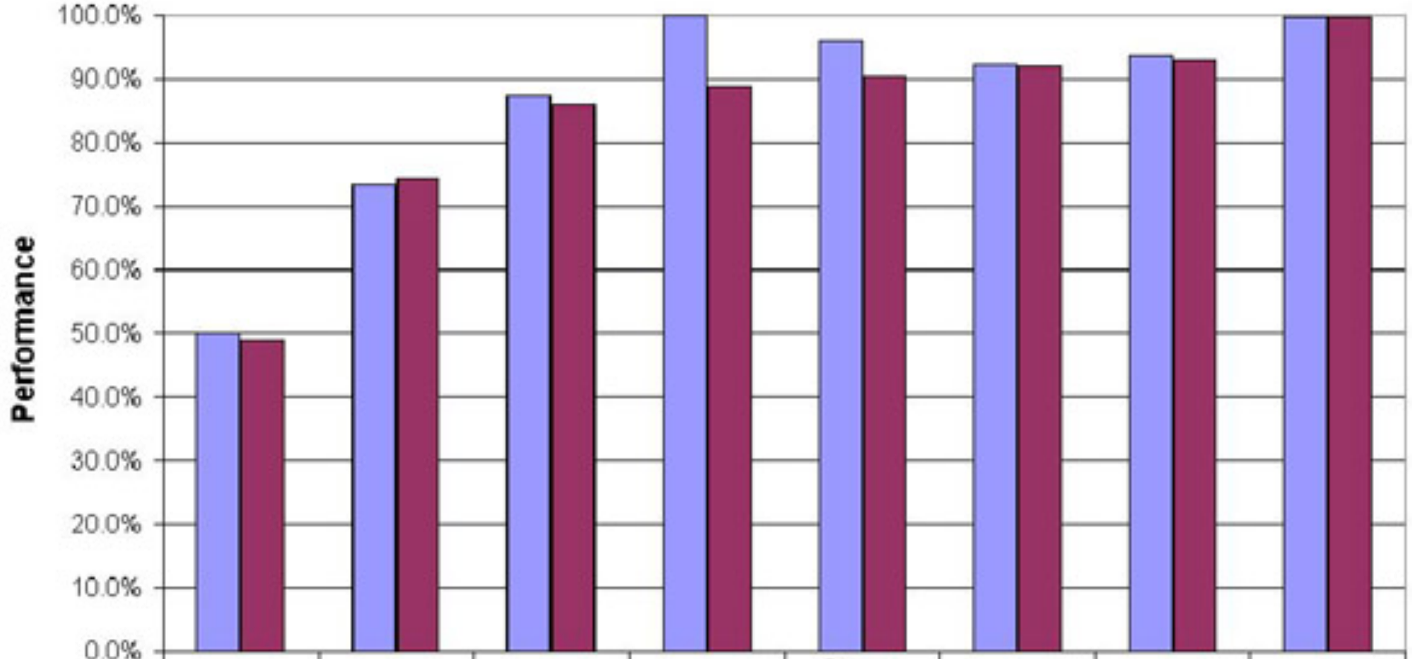
- The first 8 eigenfaces obtained from a training set of 180 male and 180 female training images (without mean removal)

A. Diaco, J. DiCarlo and J. Santos, EE368 class project, Spring 2000.

- Can be used to generate faces by adjusting 8 coefficients.
- Can be used for face recognition by nearest neighbor search in 8-d „face space."

Gender recognition using eigenfaces

Recognition accuracy using 8 eigenimages



	K-means	2nd weight	Mahalanobis	Closest match	Closest diffusion match	Linear discriminant	Local linear discriminant	Human
Training set	50.1%	73.6%	87.6%	100.0%	95.9%	92.3%	93.8%	99.8%
Test set	49.0%	74.3%	86.0%	88.8%	90.5%	92.0%	93.0%	99.8%

A. Diaco, J. DiCarlo and J. Santos, EE368 class project, Spring 2000.

Female face samples

Male face samples

Fisher linear discriminant analysis

- Eigenimage method maximizes "scatter" within the linear subspace over the entire image set – regardless of classification task

$$W_{opt} = \arg\max_{W} \left(\det\left(W R W^H \right) \right)$$

- Fisher linear discriminant analysis (1936): maximize between-class scatter, while minimizing within-class scatter

$$W_{opt} = \arg\max_{W} \left(\frac{\det\left(W R_B W^H \right)}{\det\left(W R_W W^H \right)} \right)$$

$$R_B = \sum_{i=1}^{c} N_i \left(\vec{\mu}_i - \vec{\mu} \right) \left(\vec{\mu}_i - \vec{\mu} \right)^H$$

Samples in class i

Mean in class i

$$R_W = \sum_{i=1}^{c} \sum_{\vec{\Gamma}_l \in Class(i)} \left(\vec{\Gamma}_l - \vec{\mu}_i \right) \left(\vec{\Gamma}_l - \vec{\mu}_i \right)^H$$

Fisher linear discriminant analysis (cont.)

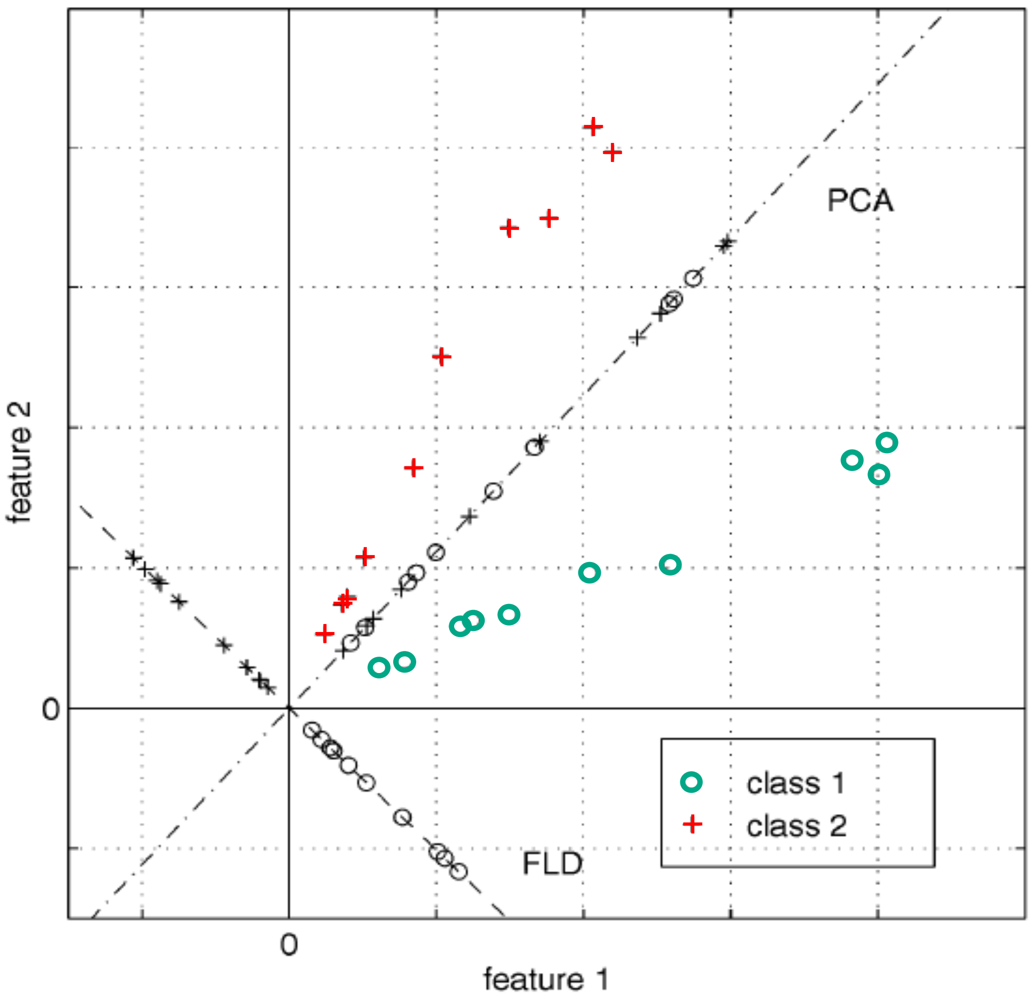
- Solution: Generalized eigenvectors $\vec{w_i}$ corresponding to the K largest eigenvalues $\{\lambda_i \mid i = 1,2,\ldots,K\}$, i.e.

$$R_B \vec{w_i} = \lambda_i R_W \vec{w_i} \;,\quad i = 1,2,\ldots,K$$

- Problem: within-class scatter matrix R_w at most of rank L-c, hence usually singular.
- Apply KLT first to reduce dimension of feature space to L-c (or less), proceed with Fisher LDA in low-dimensional space

Eigenimages vs. Fisherimages

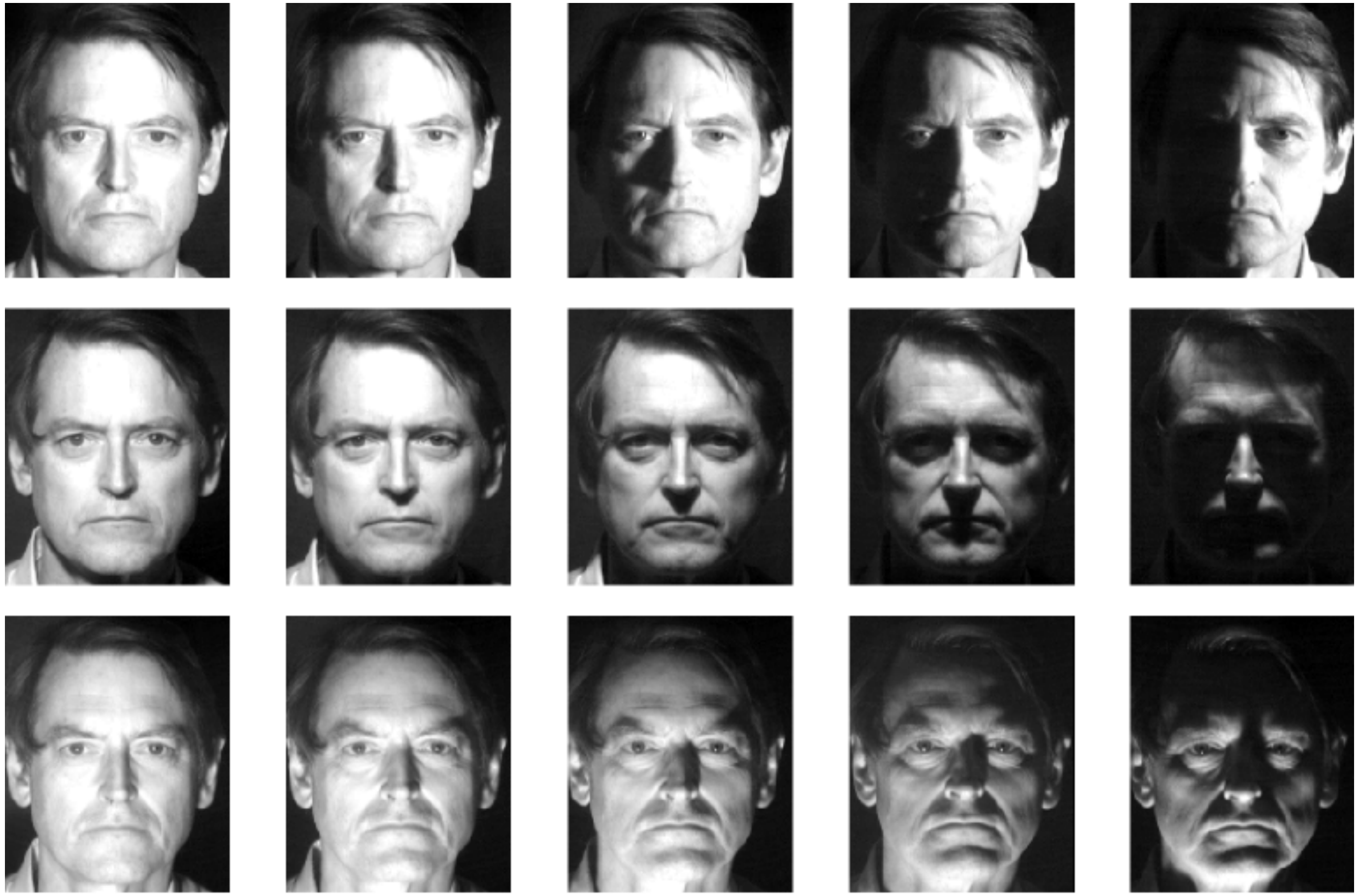
<u>2d example:</u>
Samples for 2 classes are projected onto 1d subspace using the KLT (aka PCA) or Fisher LDA (FLD). PCA preserves maximum energy, but the 2 classes are no longer distinguishable. FLD separates the classes by choosing a better 1d subspace.



[Belhumeur, Hespanha, Kriegman, 1997]

Fisherimages and varying illlumination

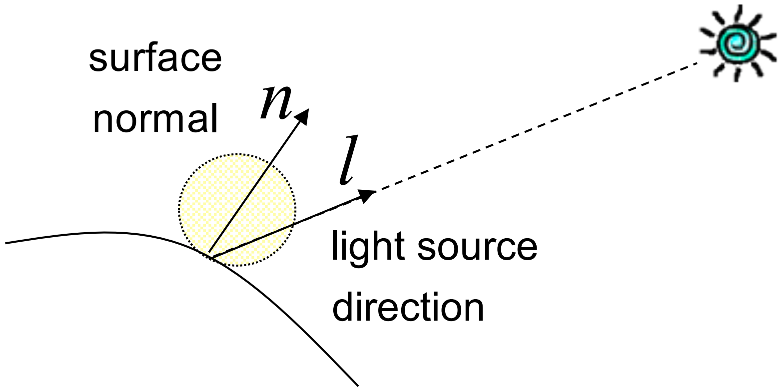
Differences due to varying illumination can be much larger than differences between faces!



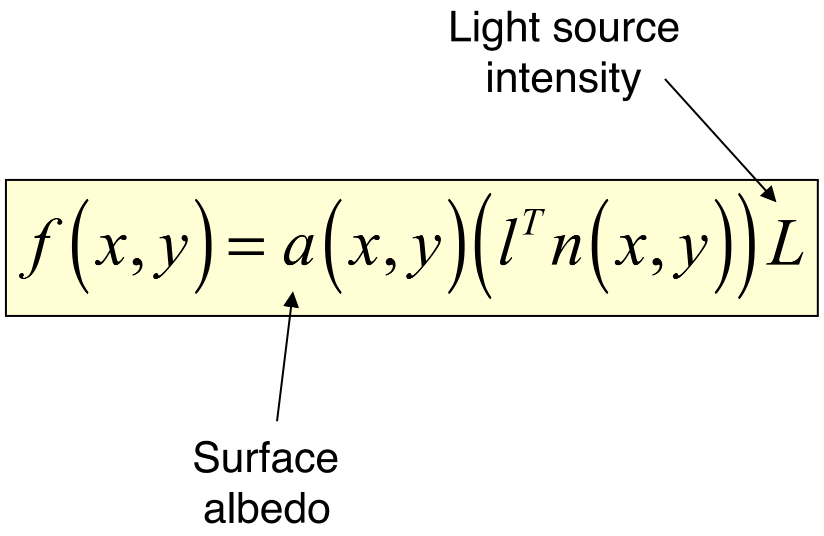
[Belhumeur, Hespanha, Kriegman, 1997]

Fisherimages and varying illlumination

- All images of same Lambertian surface with different illumination (without shadows) lie in a 3d linear subspace
- Single point source at infinity

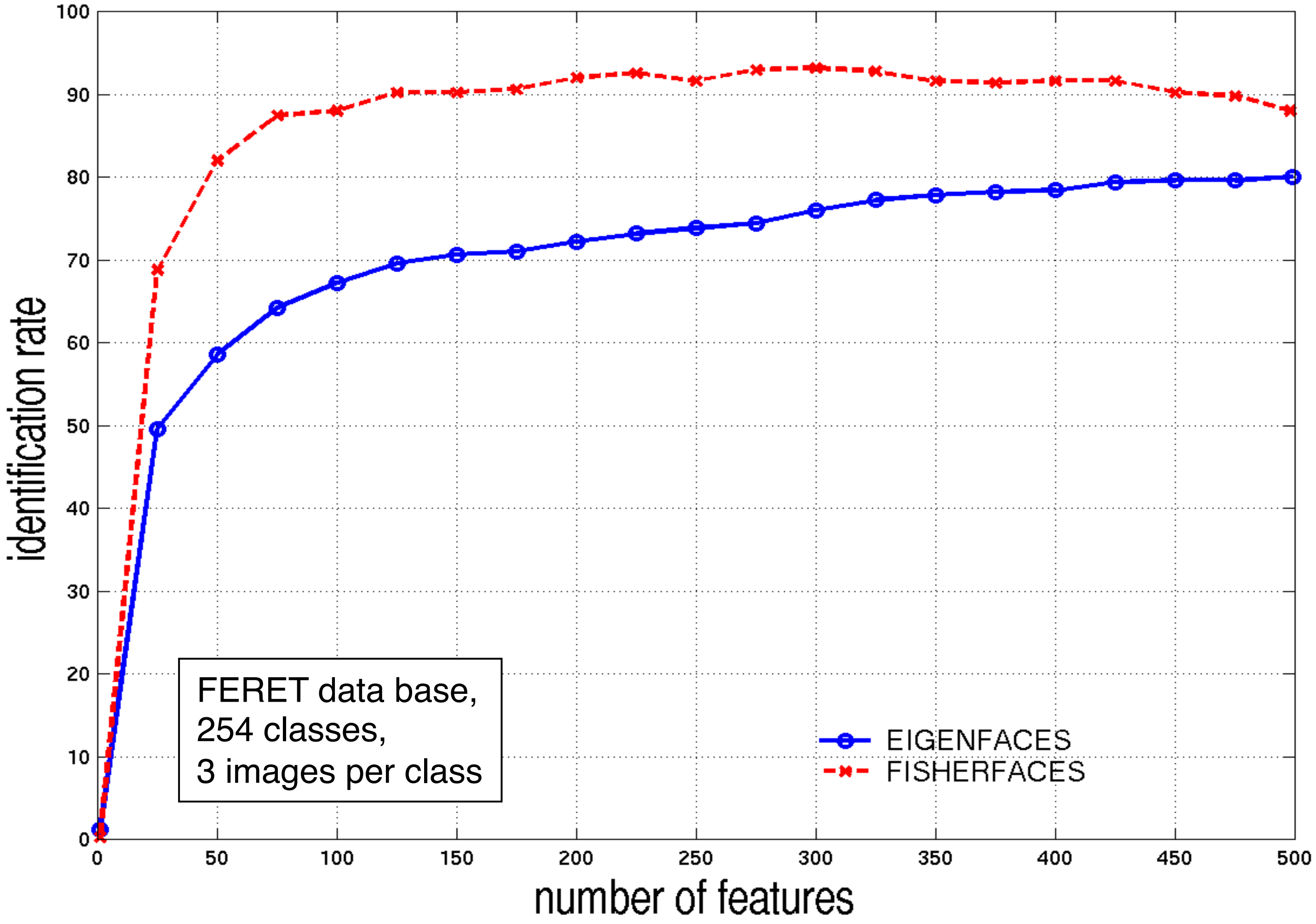


$$f(x,y) = a(x,y)\left(l^T n(x,y)\right)L$$



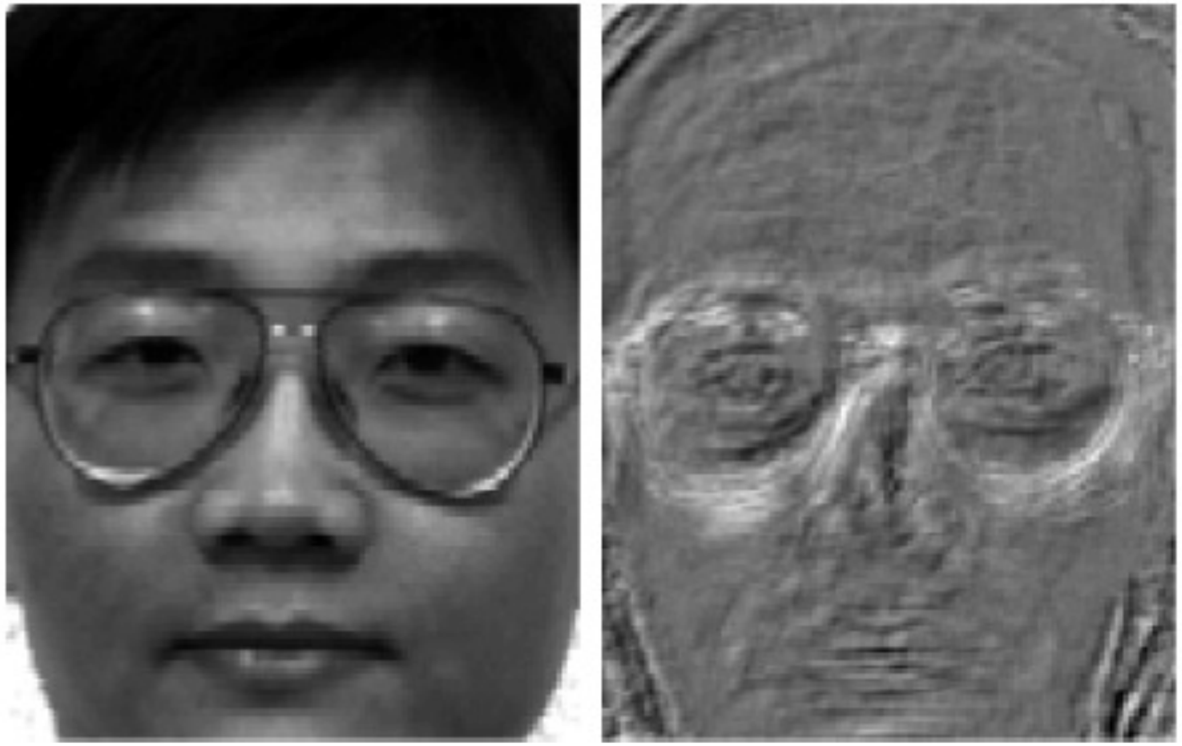
- Superposition of arbitrary number of point sources at infinity still in same 3d linear subspace, due to linear superposition of each contribution to image
- Fisherimages can eliminate within-class scatter

Face recognition with Eigenfaces and Fisherfaces



[Belhumeur, Hespanha, Kriegman, 1997]

Fisherface trained to recognize glasses



[Belhumeur, Hespanha, Kriegman, 1997]

Fisherface trained to recognize gender

Female face samples

Male face samples

Mean image

$\vec{\mu}$

Female mean

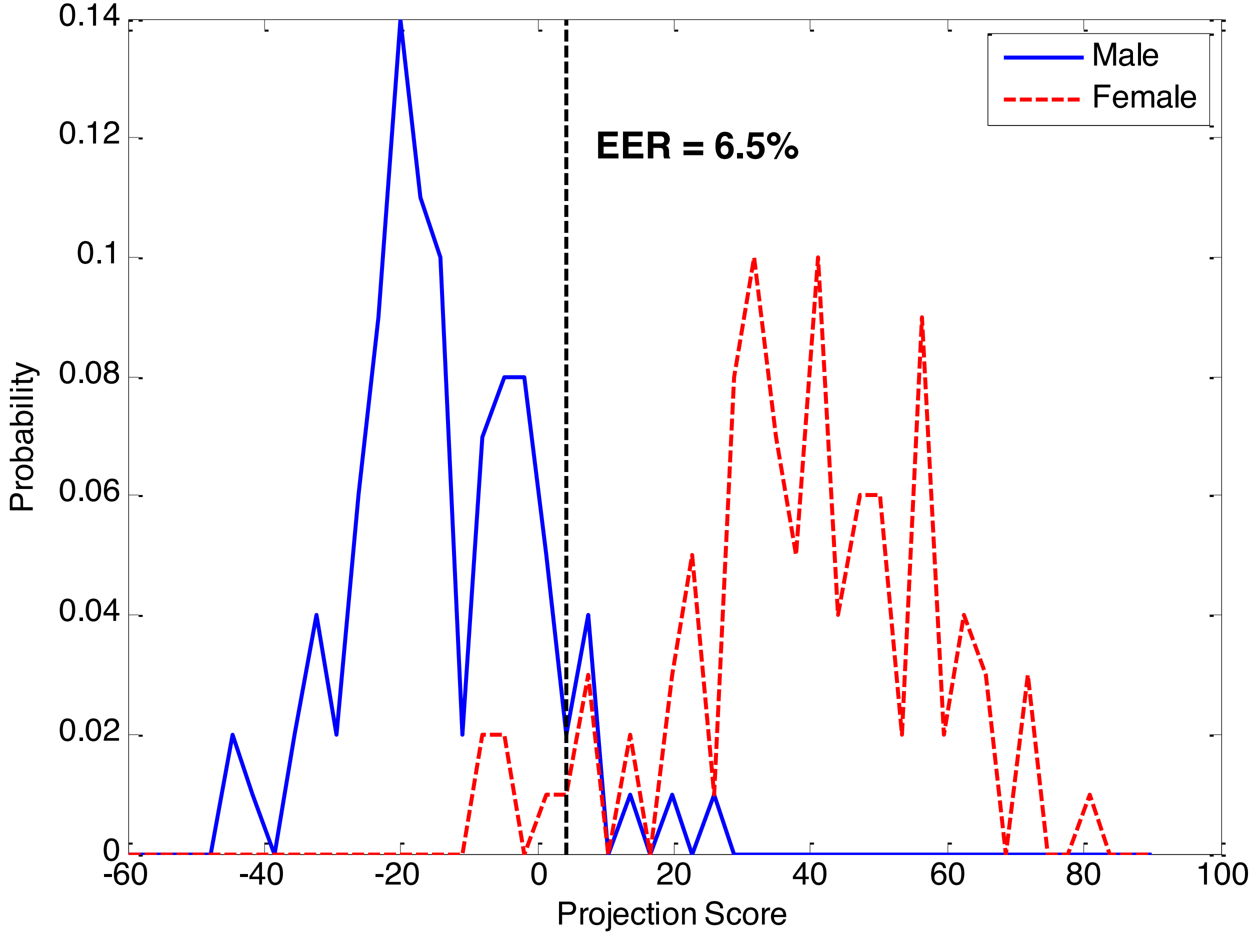
$\vec{\mu}_1$

Male mean

$\vec{\mu}_2$

Fisherface

Gender recognition using fisherface



Gender recognition using 2nd eigenface