Lecture 17: Face Recognition

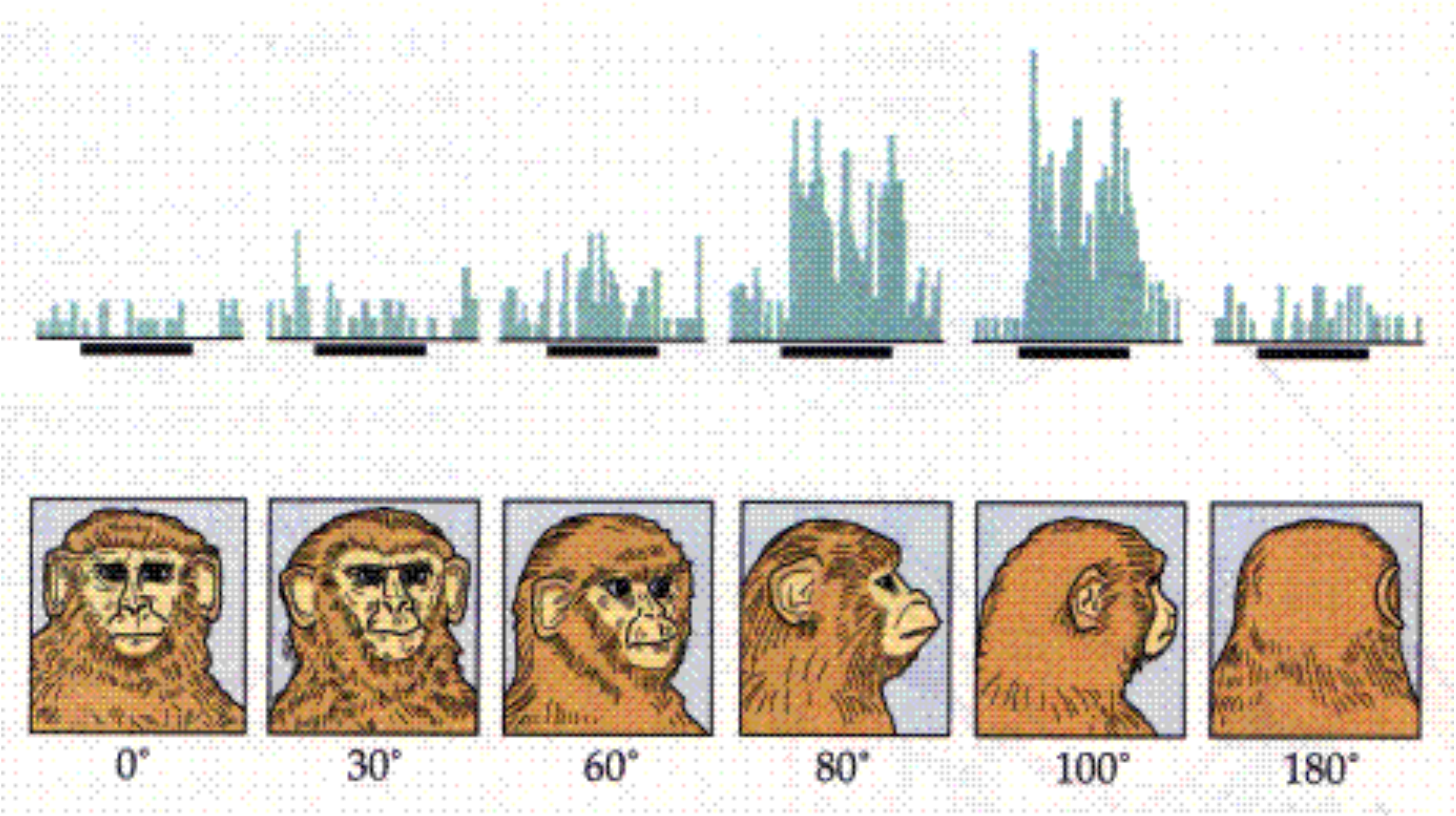
Professor Fei-Fei Li
Stanford Vision Lab

What we will learn today

1. Turk and Pentland, Eigenfaces for Recognition, *Journal of Cognitive Neuroscience* **3** (1): 71–86.

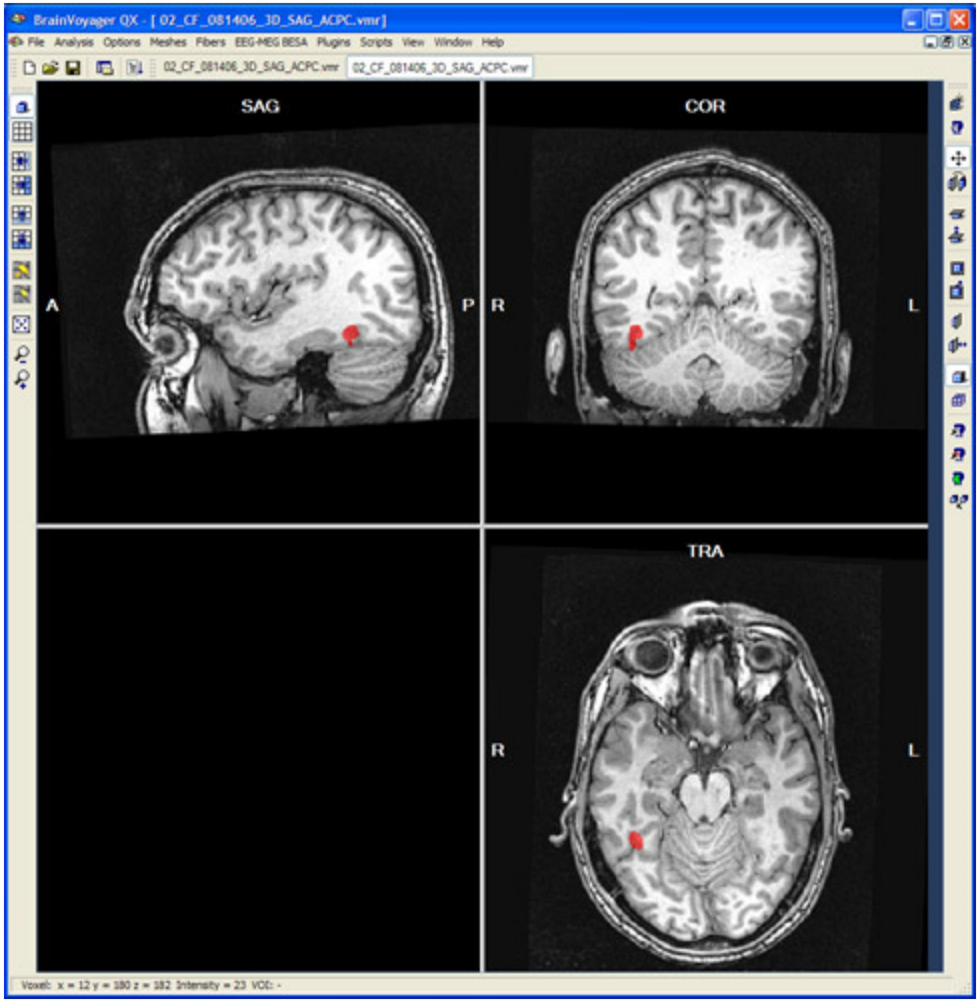
2. P. Belhumeur, J. Hespanha, and D. Kriegman. "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection". *IEEE Transactions on pattern analysis and machine intelligence* **19** (7): 711. 1997.

“Faces” in the brain



Courtesy of Johannes M. Zanker

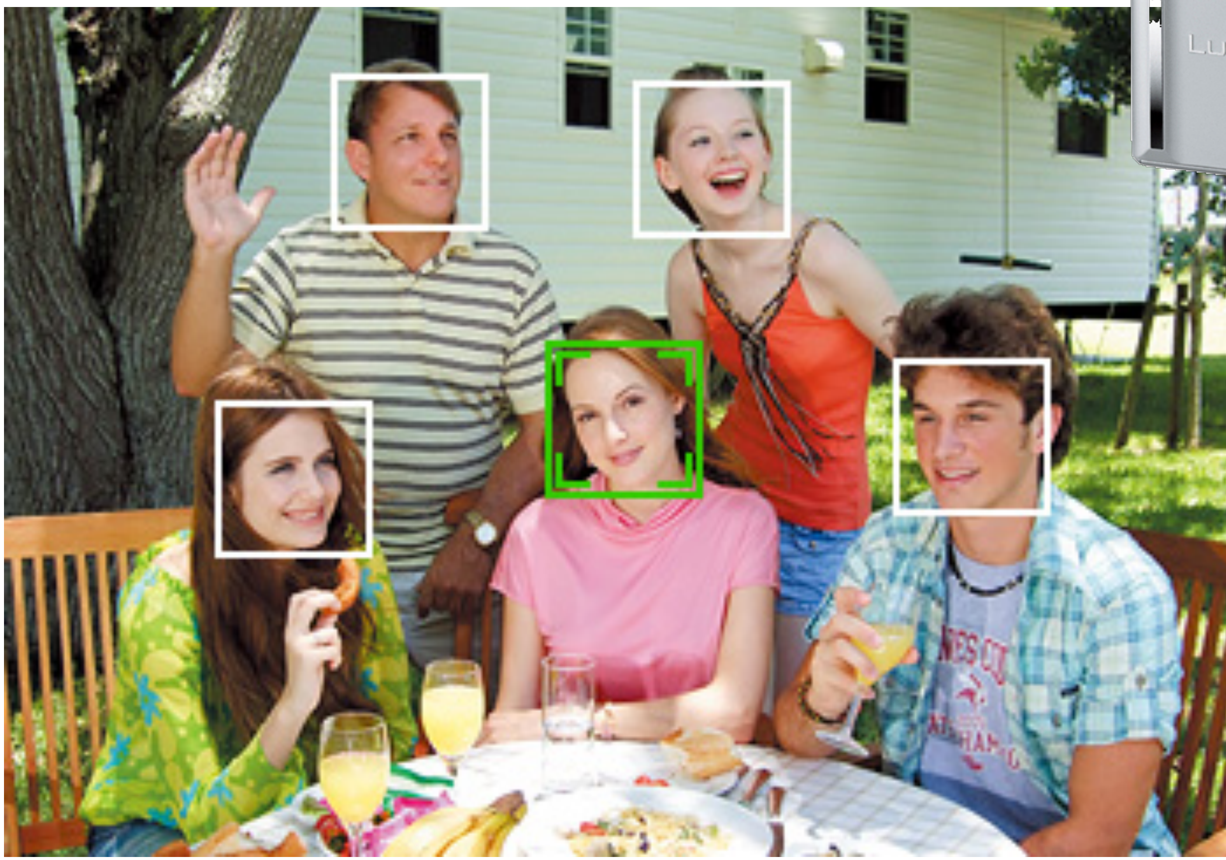
"Faces" in the brain



Kanwisher, et al. 1997

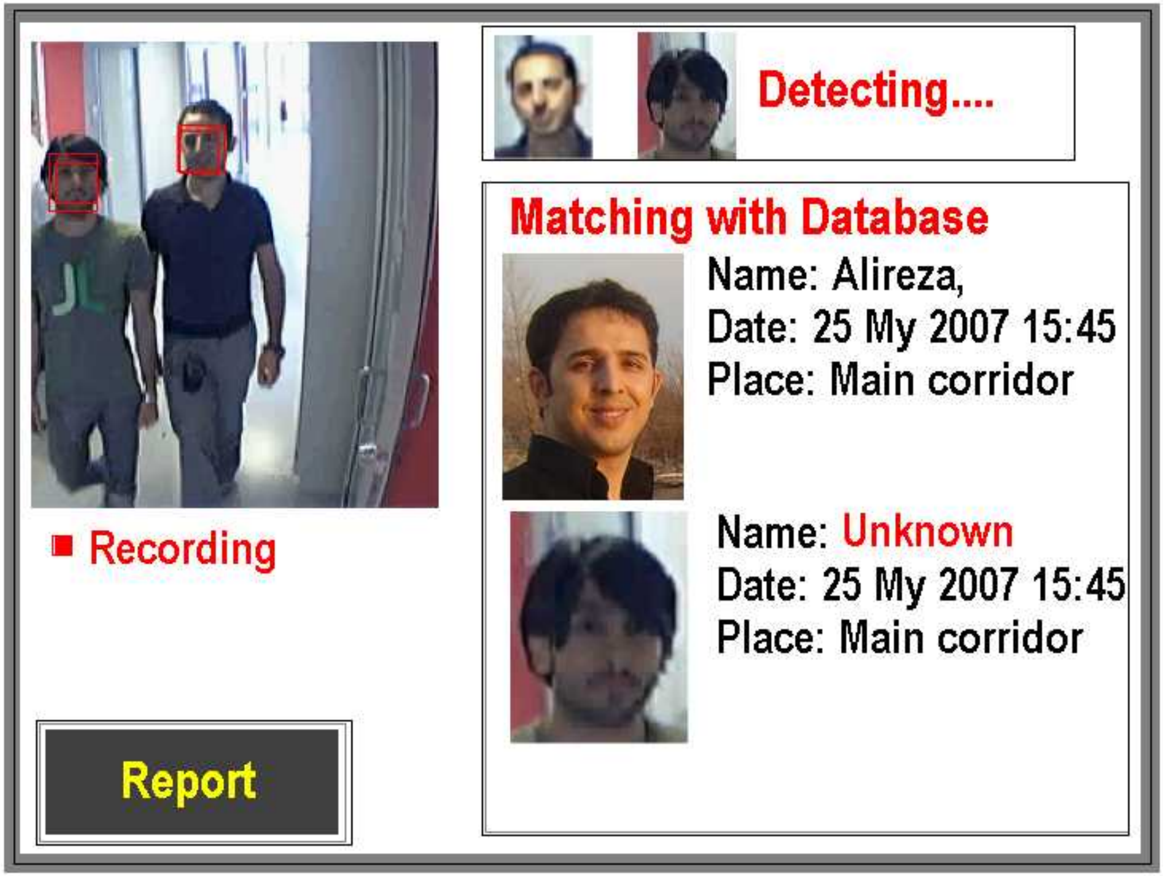
Face Recognition

- Digital photography



Face Recognition

- Digital photography
- Surveillance

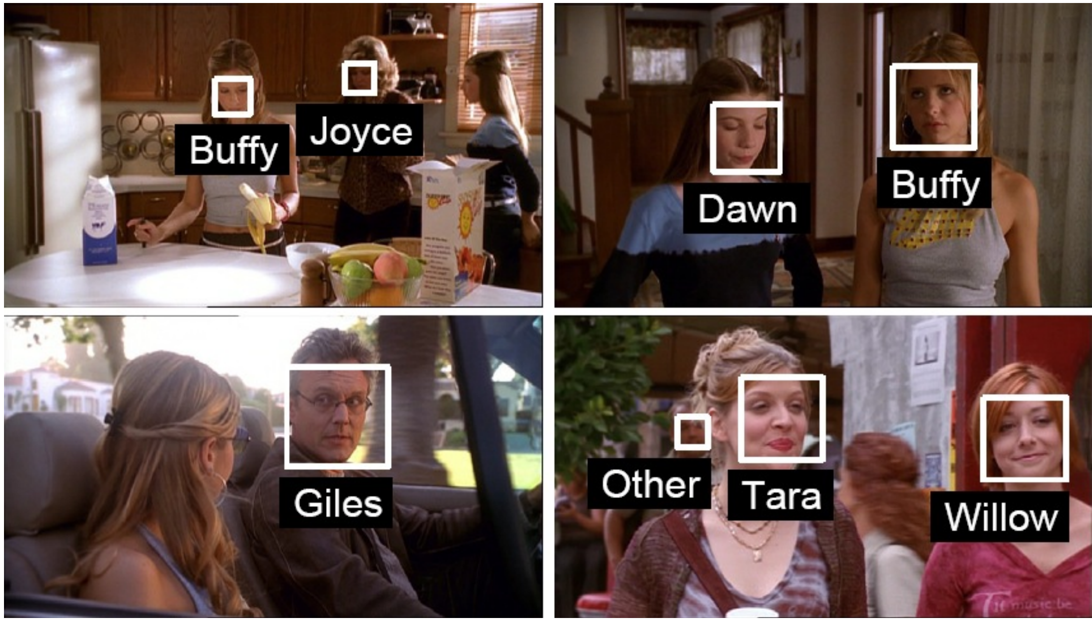


Face Recognition

- Digital photography
- Surveillance
- Album organization

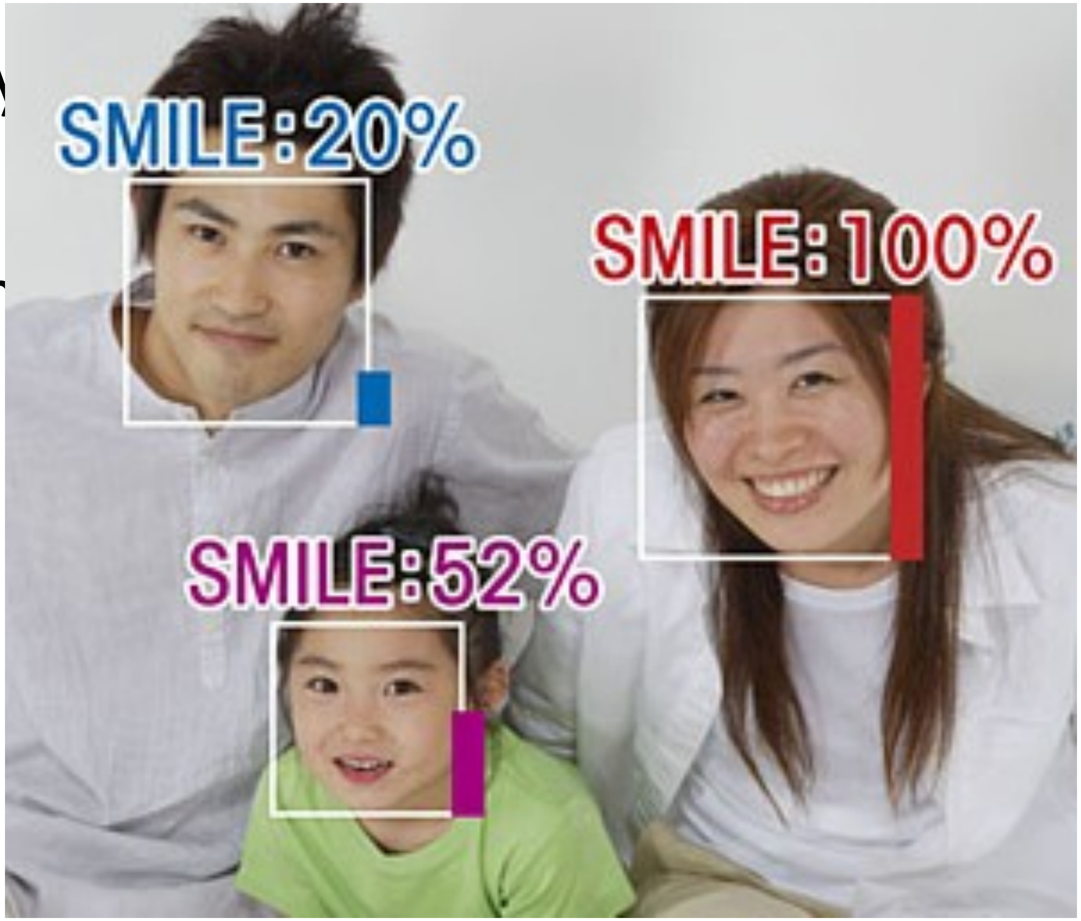
Face Recognition

- Digital photography
- Surveillance
- Album organization
- Person tracking/id.



Face Recognition

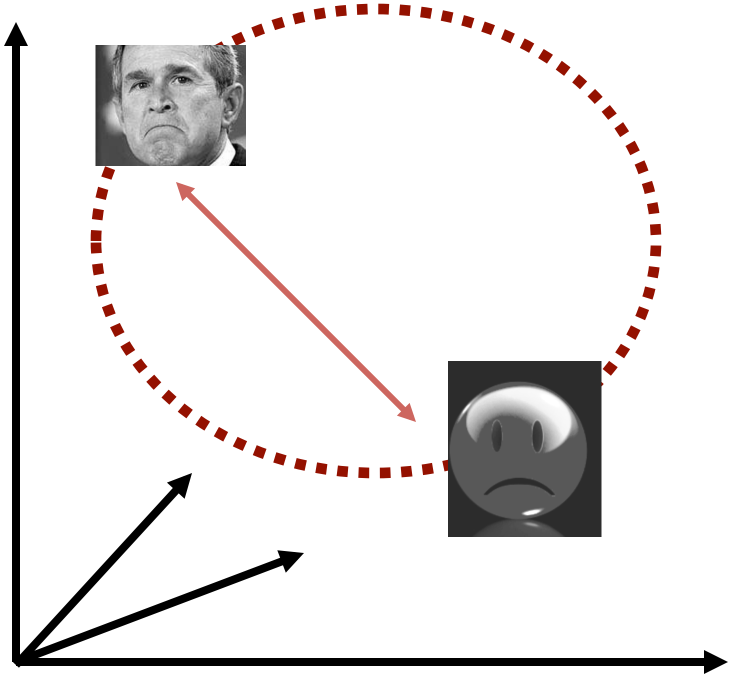
- Digital photography
- Surveillance
- Album organization
- Person tracking/id.
- Emotions and expressions



Face Recognition

- Digital photography
- Surveillance
- Album organization
- Person tracking/id.
- Emotions and expressions
- Security/warfare
- Tele-conferencing
- Etc.

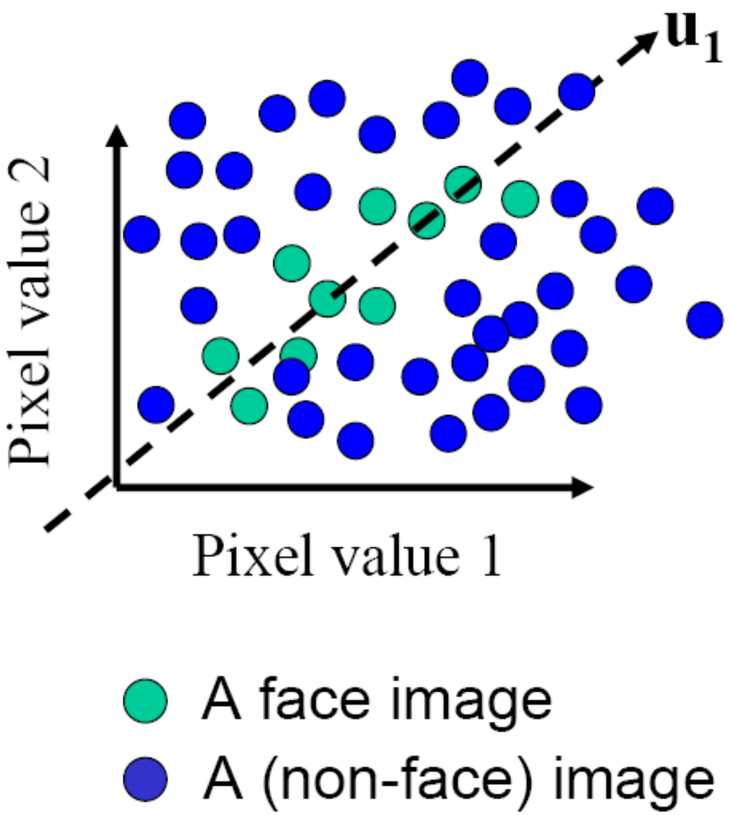
The Space of Faces



- An image is a point in a high dimensional space
  - If represented in grayscale intensity, an N x M image is a point in R^{NM}
  - E.g. 100x100 image = 10,000 dim

Slide credit: Chuck Dyer, Steve Seitz, Nishino

The Space of Faces



- An image is a point in a high dimensional space
  - If represented in grayscale intensity, an N x M image is a point in R^{NM}
  - E.g. 100x100 image = 10,000 dim
- However, relatively few high dimensional vectors correspond to valid face images
- We want to effectively model the subspace of face images

Slide credit: Chuck Dyer, Steve Seitz, Nishino

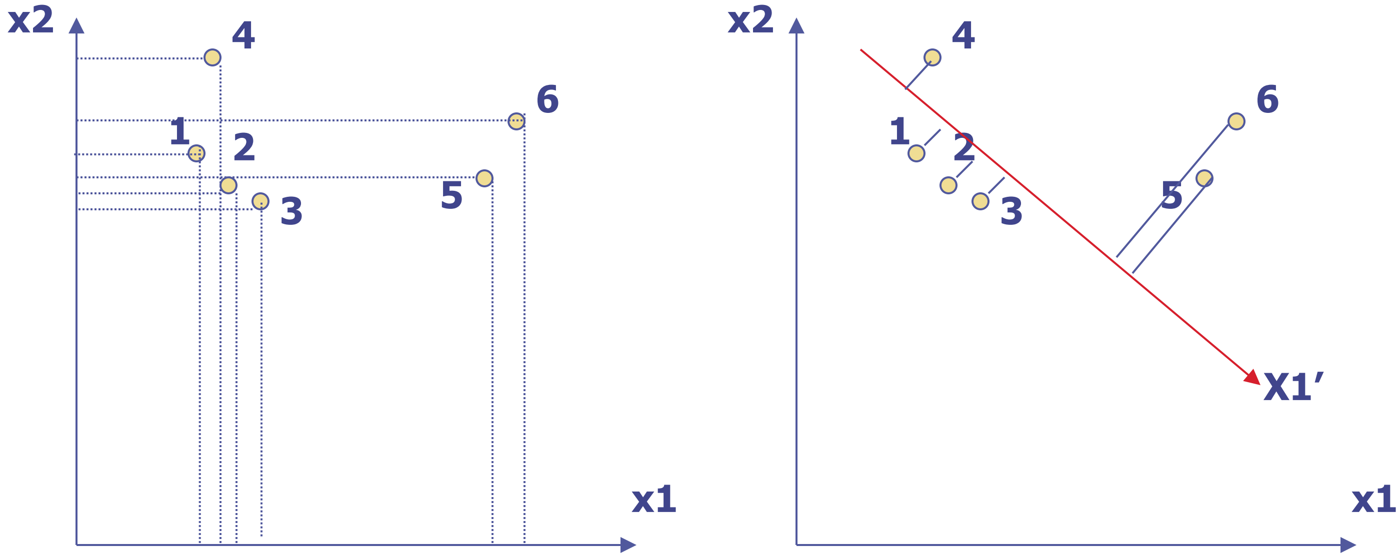
Key Idea

- So, compress them to a low-dimensional subspace that captures key appearance characteristics of the visual DOFs.

- USE PCA for estimating the sub-space (dimensionality reduction)

- Compare two faces by projecting the images into the subspace and measuring the EUCLIDEAN distance between them.

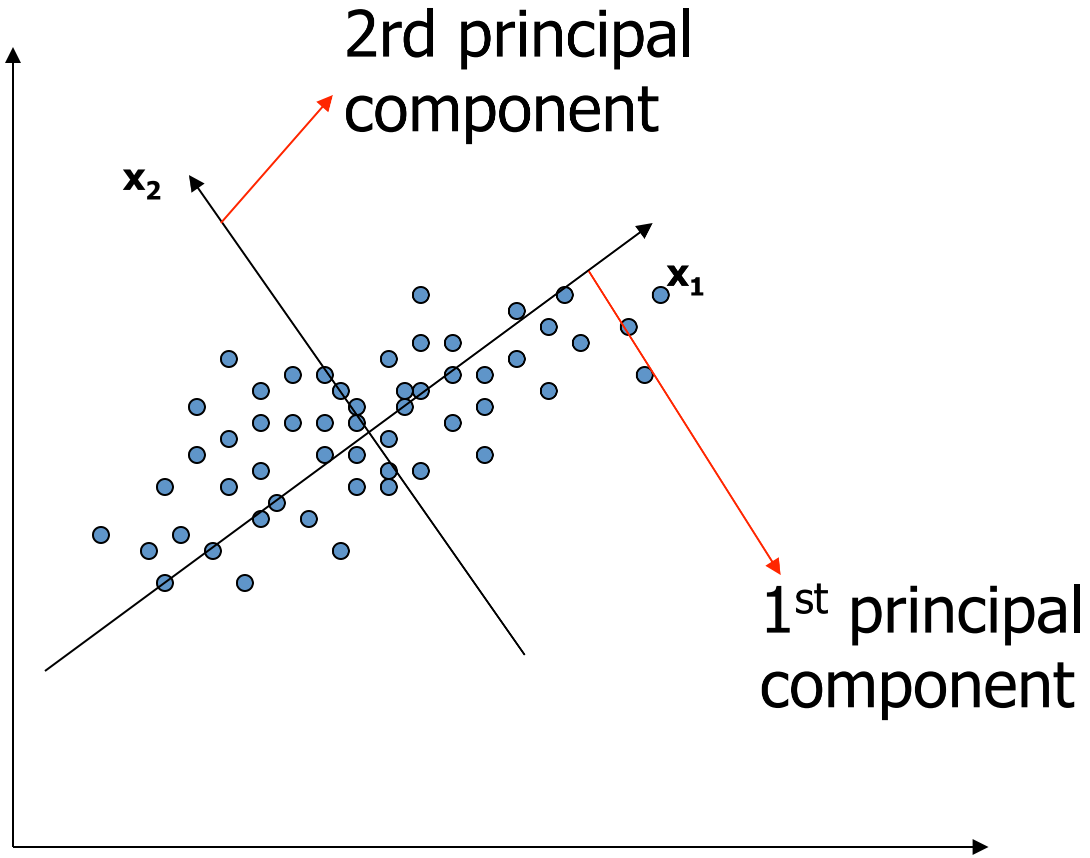
USE PCA for estimating the sub-space

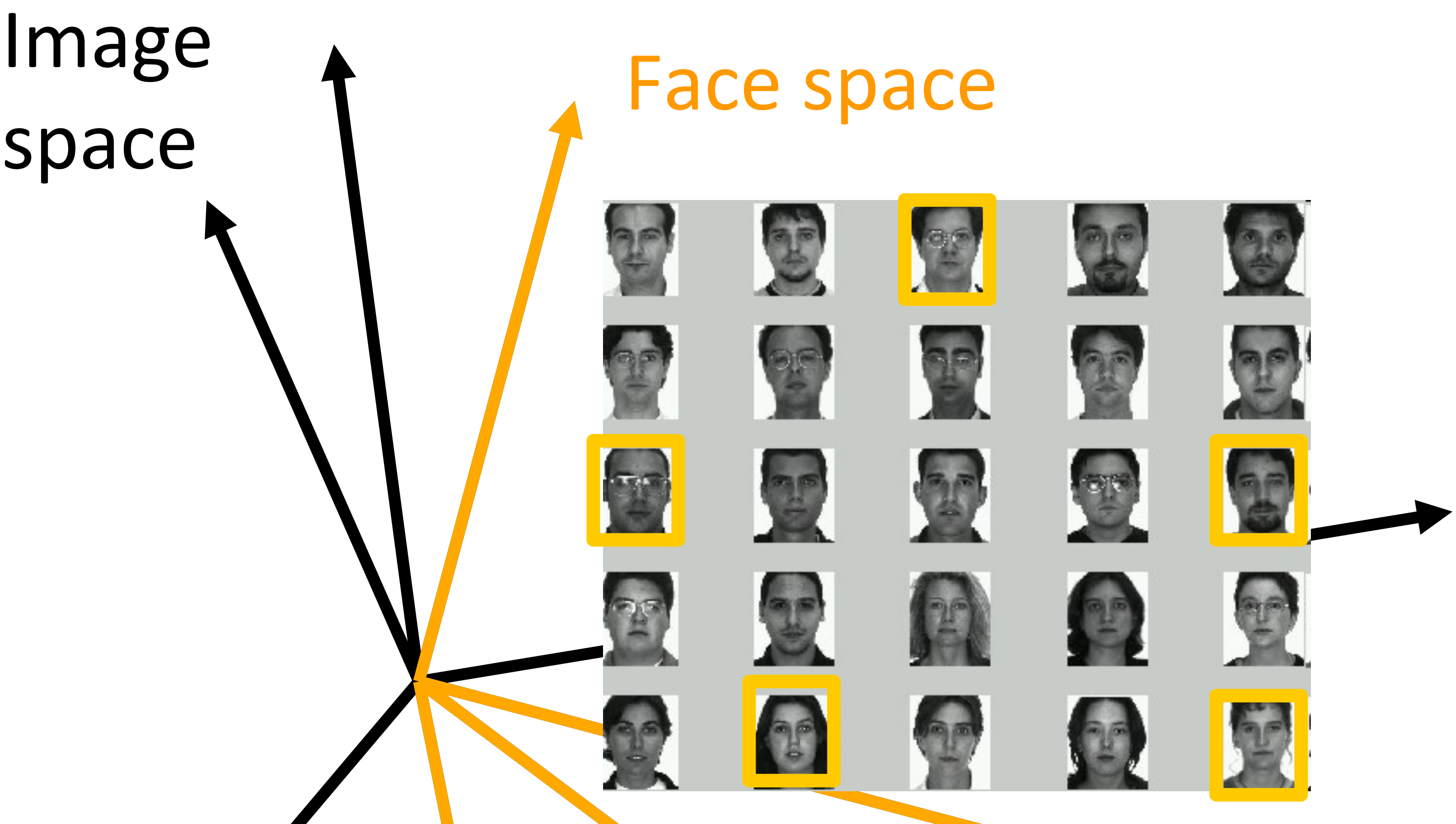


PCA projection

- Computes n-dim subspace such that the projection of the data points onto the subspace has the largest variance among all n-dim subspaces.

USE PCA for estimating the sub-space





- Computes n-dim subspace such that the projection of the data points onto the subspace has the largest variance among all n-dim subspaces.
- Maximize the scatter of the training images in face space

PCA: Mathematical formulation

PCA = eigenvalue decomposition of a data covariance matrix

Define a transformation, W,

$$y_j = W^T x_j \qquad j = 1, 2 \ldots N$$

- y_j → **m-dimensional**
- W → **Orthonorm** $W \in \mathbb{R}^{n \times m}$
- x_j → **n-dimensional**

$$S_T = \sum_{j=1}^{N} (x_j - \bar{x})(x_j - \bar{x})^T$$ = Data Scatter matrix

$$\widetilde{S}_T = \sum_{j=1}^{N} (y_j - \bar{y})(y_j - \bar{y})^T = W^T S_T W$$ = Transf. data scatter matrix

$$W_{opt} = \arg\max_W \left| W^T S_T W \right| = \underbrace{\left[v_1 \; v_2 \cdots v_m \right]}_{\text{Eigenvectors of } S_T}$$

Eigenfaces: key idea

- Assume that most face images lie on a low-dimensional subspace determined by the first k (k<<d) directions of maximum variance
- Use PCA to determine the vectors or "eigenfaces" that span that subspace
- Represent all face images in the dataset as linear combinations of eigenfaces

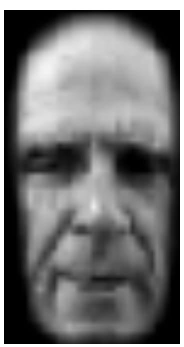
M. Turk and A. Pentland, Face Recognition using Eigenfaces, CVPR 1991

Projecting onto the Eigenfaces

- The eigenfaces $\mathbf{v_1}, \ldots, \mathbf{v_K}$ span the space of faces
  - A face is converted to eigenface coordinates by

$$\mathbf{x} \rightarrow (\underbrace{(\mathbf{x}-\overline{\mathbf{x}}) \cdot \mathbf{v}_1}_{a_1}, \underbrace{(\mathbf{x}-\overline{\mathbf{x}}) \cdot \mathbf{v}_2}_{a_2}, \ldots, \underbrace{(\mathbf{x}-\overline{\mathbf{x}}) \cdot \mathbf{v}_K}_{a_K})$$

$$\mathbf{x} \approx \overline{\mathbf{x}} + a_1\mathbf{v_1} + a_2\mathbf{v_2} + \ldots + a_K\mathbf{v_K}$$



$\mathbf{x}$

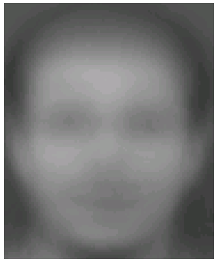
$a_1\mathbf{v}_1$ $a_2\mathbf{v}_2$ $a_3\mathbf{v}_3$ $a_4\mathbf{v}_4$ $a_5\mathbf{v}_5$ $a_6\mathbf{v}_6$ $a_7\mathbf{v}_7$ $a_8\mathbf{v}_8$

Algorithm

Training

1. Align training images $x_1, x_2, \ldots, x_N$

Note that each image is formulated into a long vector!



2. Compute average face $u = 1/N \sum x_i$

3. Compute the difference image $\varphi_i = x_i - u$

Algorithm

4. Compute the covariance matrix (total scatter matrix)

$$S_T = (1/N) \Sigma \phi_i \phi_i^T = BB^T, \quad B=[\phi_1, \phi_2 \dots \phi_N]$$

5. Compute the eigenvectors of the covariance matrix S_T
6. Compute training projections $a_1, a_2 \dots a_N$

Testing

1. Take query image X
2. Project X into Eigenface space (W = {eigenfaces}) and compute projection $\omega_i = W(X - u)$,
3. Compare projection ω_i with all training N projections a_i

Illustration of Eigenfaces

- The visualization of eigenvectors:

These are the first 4 eigenvectors from a training set of 400 images (ORL Face Database).

Eigenfaces look somewhat like generic faces.

Reconstruction and Errors

P = 4

P = 200

P = 400

- Only selecting the top P eigenfaces → reduces the dimensionality.
- Fewer eigenfaces result in more information loss, and hence less discrimination between faces.

Summary for Eigenface

Pros

- Non-iterative, globally optimal solution

Limitations

- PCA projection is **optimal for reconstruction** from a low dimensional basis, but **may NOT be optimal for discrimination…**
  - See supplementary materials for "Linear Discriminative Analysis", aka "Fisherfaces"

What we have learned today

1. Turk and Pentland, Eigenfaces for Recognition, *Journal of Cognitive Neuroscience* **3** (1): 71–86.

2. P. Belhumeur, J. Hespanha, and D. Kriegman. "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection". *IEEE Transactions on pattern analysis and machine intelligence* **19** (7): 711. 1997.

Supplementary materials

1. Turk and Pentland, Eigenfaces for Recognition, *Journal of Cognitive Neuroscience* **3** (1): 71–86.

2. P. Belhumeur, J. Hespanha, and D. Kriegman. "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection". *IEEE Transactions on pattern analysis and machine intelligence* **19** (7): 711. 1997.

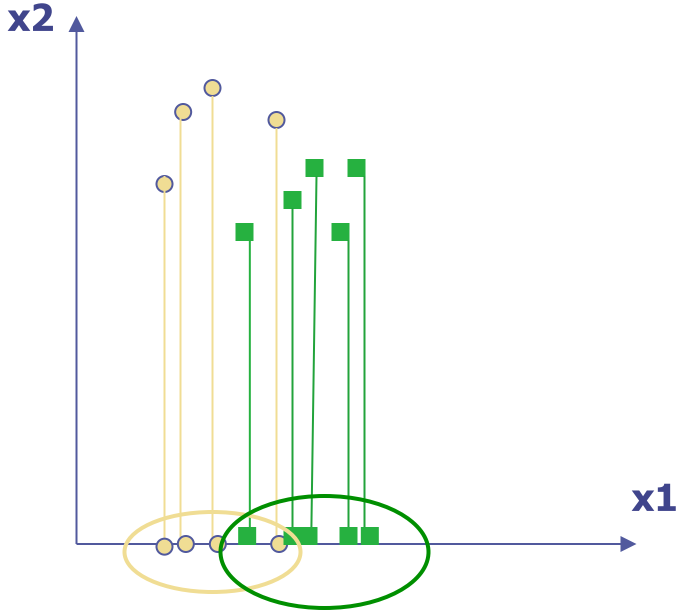
Fisher Faces:
Linear Discriminant Analysis

Linear Discriminant Analysis (LDA)
Fisher's Linear Discriminant (FLD)

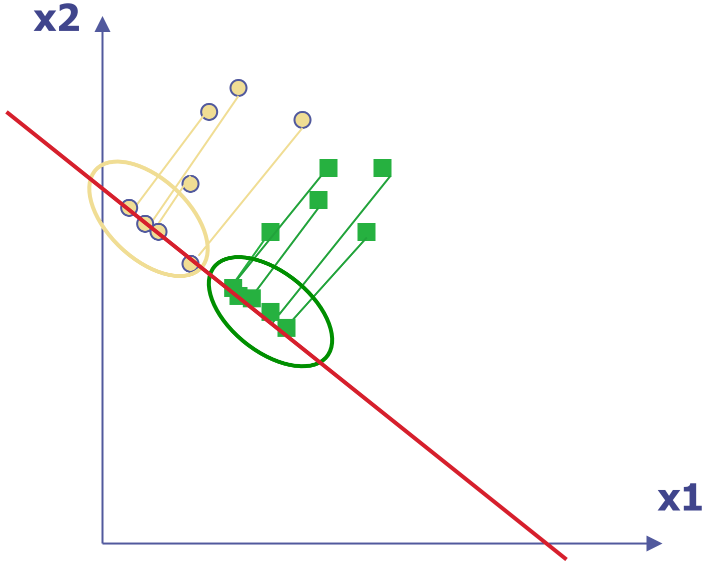
- Eigenfaces exploit the max scatter of the training images in face space
- Fisherfaces attempt to maximise the **between class scatter**, while minimising the **within class scatter**.

Illustration of the Projection

- Using two classes as example:

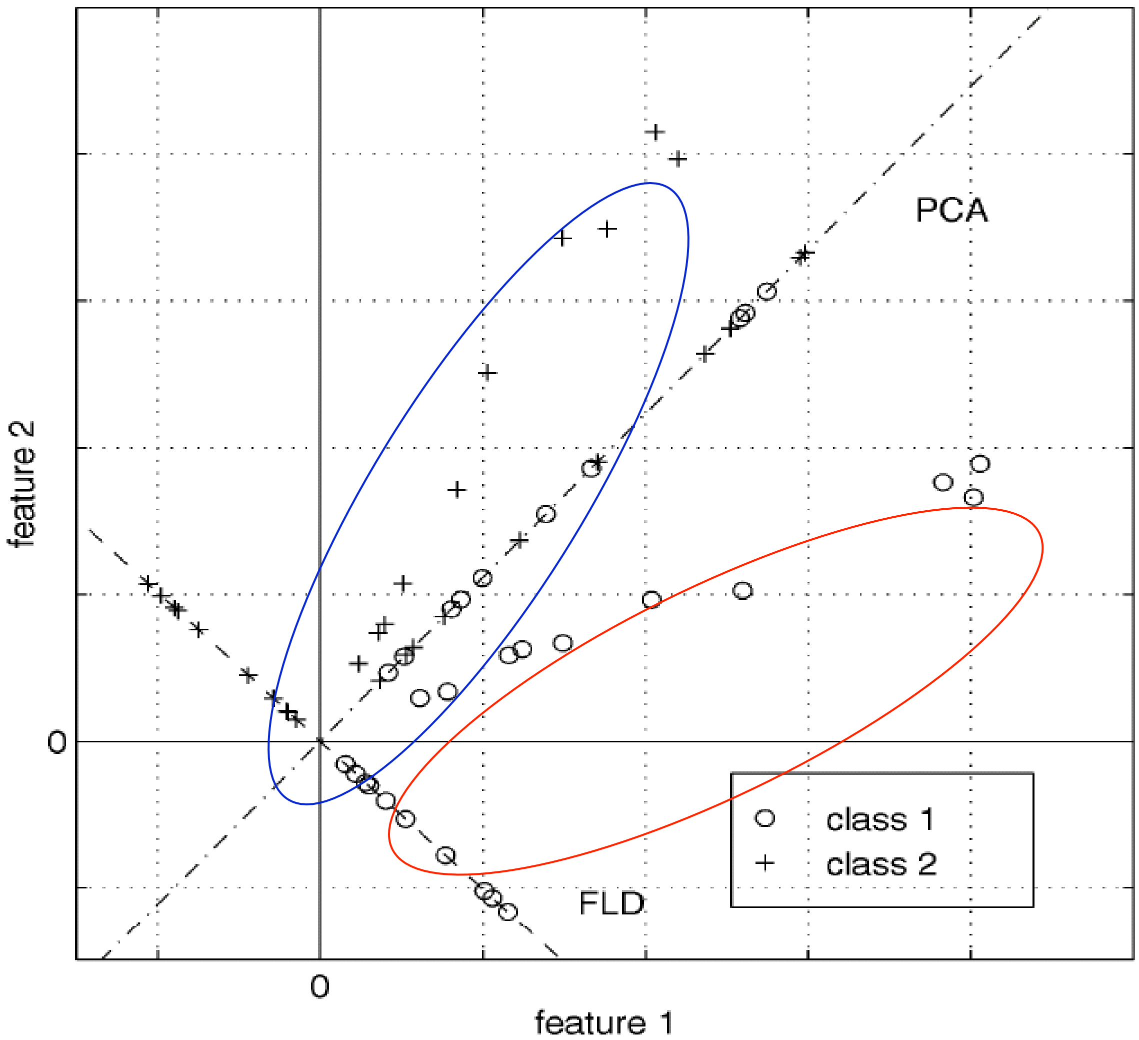


Poor Projection



Good

Comparing with PCA



Variables

- N Sample images: $\{x_1, \cdots, x_N\}$
- c classes: $\{\chi_1, \cdots, \chi_c\}$
- Average of each class: $\mu_i = \frac{1}{N_i} \sum_{x_k \in \chi_i} x_k$
- Total average: $\mu = \frac{1}{N} \sum_{k=1}^{N} x_k$

Scatters

- Scatter of class i:

$$S_i = \sum_{x_k \in \chi_i} (x_k - \mu_i)(x_k - \mu_i)^T$$

- Within class scatter:

$$S_W = \sum_{i=1}^{c} S_i$$

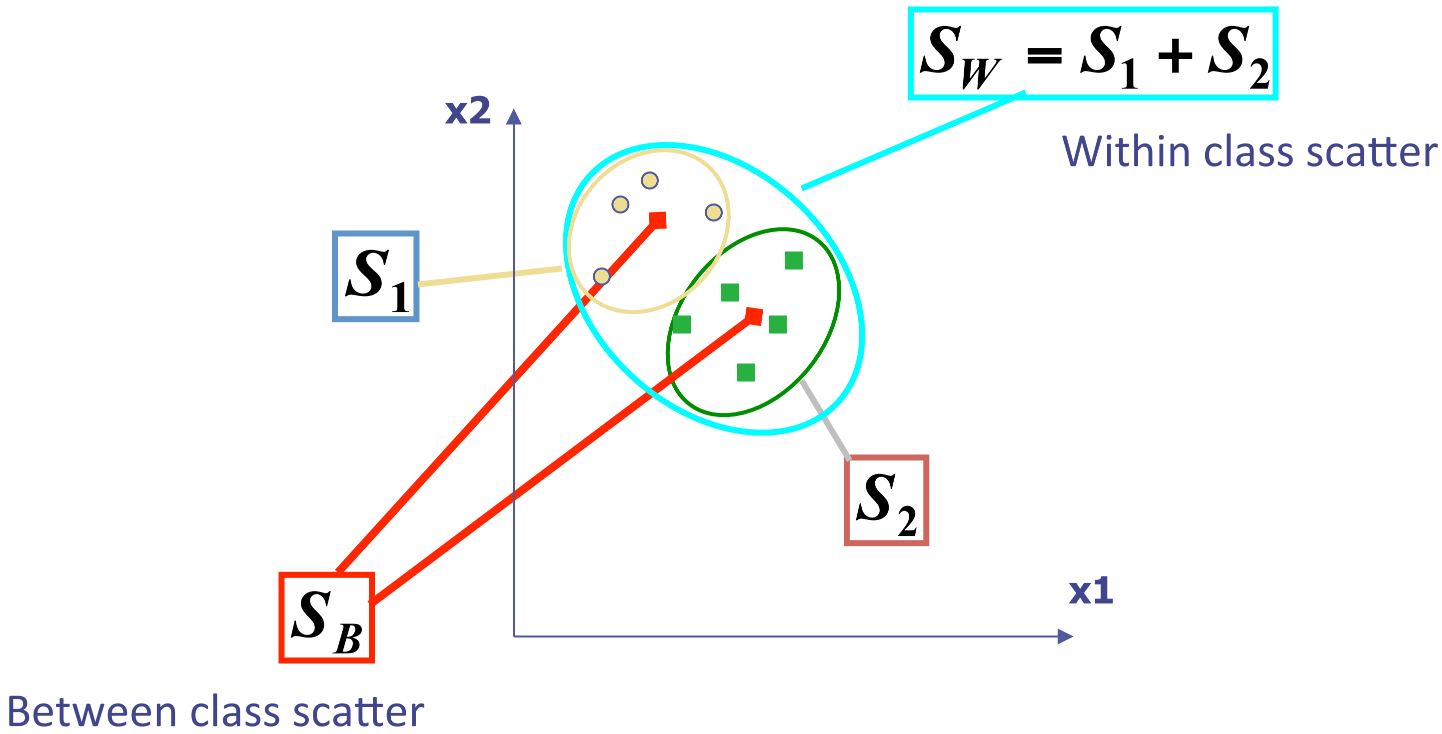
- Between class scatter:

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

- Total scatter:

$$S_T = S_W + S_B$$

Illustration



$$S_i = \sum_{x_k \in \chi_i} (x_k - \mu_i)(x_k - \mu_i)^T$$

$$S_W = \sum_{i=1}^{c} S_i$$

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

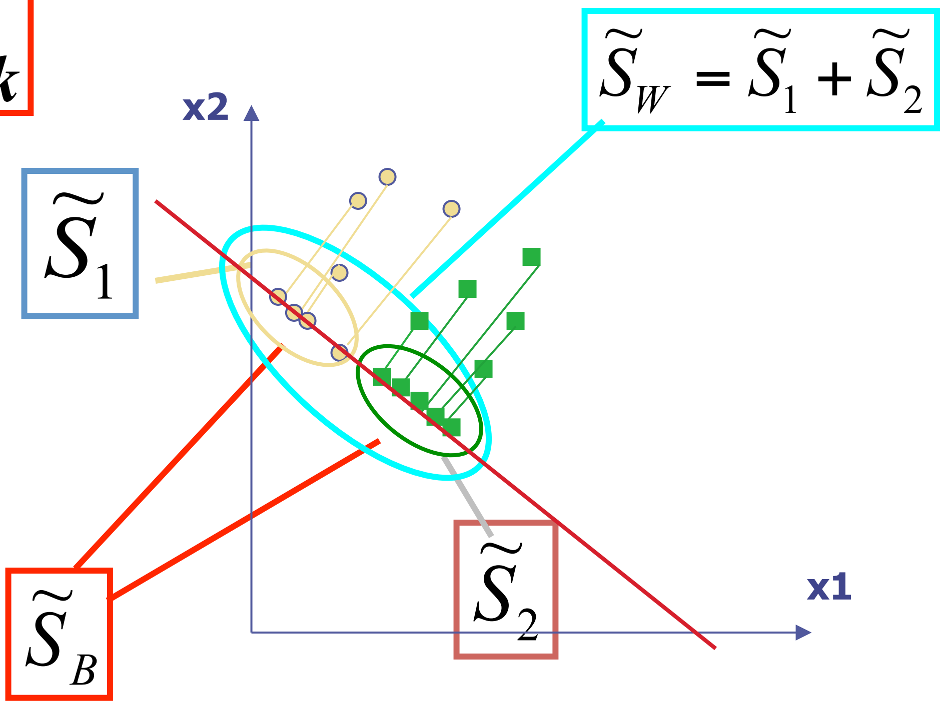
Mathematical Formulation (1)

- After projection: $y_k = W^T x_k$

- Between class scatter (of y's): $\tilde{S}_B = W^T S_B W$
- Within class scatter (of y's): $\tilde{S}_W = W^T S_W W$

Illustration

$$y_k = W^T x_k$$



$$S_W = \sum_{i=1}^{c} S_i$$

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

$$\tilde{S}_W = W^T S_W W$$

$$\tilde{S}_B = W^T S_B W$$

Mathematical Formulation

- The desired projection:

$$W_{opt} = \arg\max_{\mathbf{W}} \frac{\left|\widetilde{S}_B\right|}{\left|\widetilde{S}_W\right|} = \arg\max_{\mathbf{W}} \frac{\left|W^T S_B W\right|}{\left|W^T S_W W\right|}$$

- How is it found ? → Generalized Eigenvectors

$$S_B w_i = \lambda_i S_W w_i \qquad i = 1, \ldots, m$$

- If S_w has full rank, the generalized eigenvectors are eigenvectors of $S_W^{-1} S_B$ with largest eigen-values

Training/ Testing

Projection in Eigenface

Projection $\omega_i = W_{opt} (X - u)$,

W_{opt} = {fisher-faces}

Results: Eigenface vs. Fisherface (1)

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

- Variation in Facial Expression, Eyewear, and Lighting

Eigenface vs. Fisherface (2)