



# *EIGENFACES*

HYUN SOO PARK

# *CHALLENGES OF VISUAL RECOGNITION*



# *CHALLENGES OF VISUAL RECOGNITION*



# CHALLENGES OF VISUAL RECOGNITION

- Appearance
  - DOF: texture, illumination, material, shading, ...
- Shape
  - DOF: object category, geometric pose, viewpoint, ...



# CHALLENGES OF VISUAL RECOGNITION

- **Appearance**
  - DOF: texture, illumination, material, shading, ...
- **Shape**
  - DOF: object category, geometric pose, viewpoint, ...



# *SPACE OF APPEARANCE (FIXED SHAPE)*



$$x \in \mathbb{R}^D$$

Template

High dimension (D)

e.g., D: 10,000 = 100 x 100

# SPACE OF APPEARANCE (FIXED SHAPE)



Template

$$x \in \mathbb{R}^D$$

High dimension (D)

e.g.,  $D: 10,000 = 100 \times 100$

Naïve face detection algorithm:



$x$



$y$

Use NCC or SSD to measure similarity.

maximize  $corr(x, y)$

minimize  $\|x - y\|^2$

Why not working?

# *SPACE OF FACE APPEARANCE*





# *SPACE OF FACE APPEARANCE*



# ***MISS KOREA CONTESTANTS***

Observation: not all pixels are equally informative to detect a face



# *MISS KOREA CONTESTANTS*

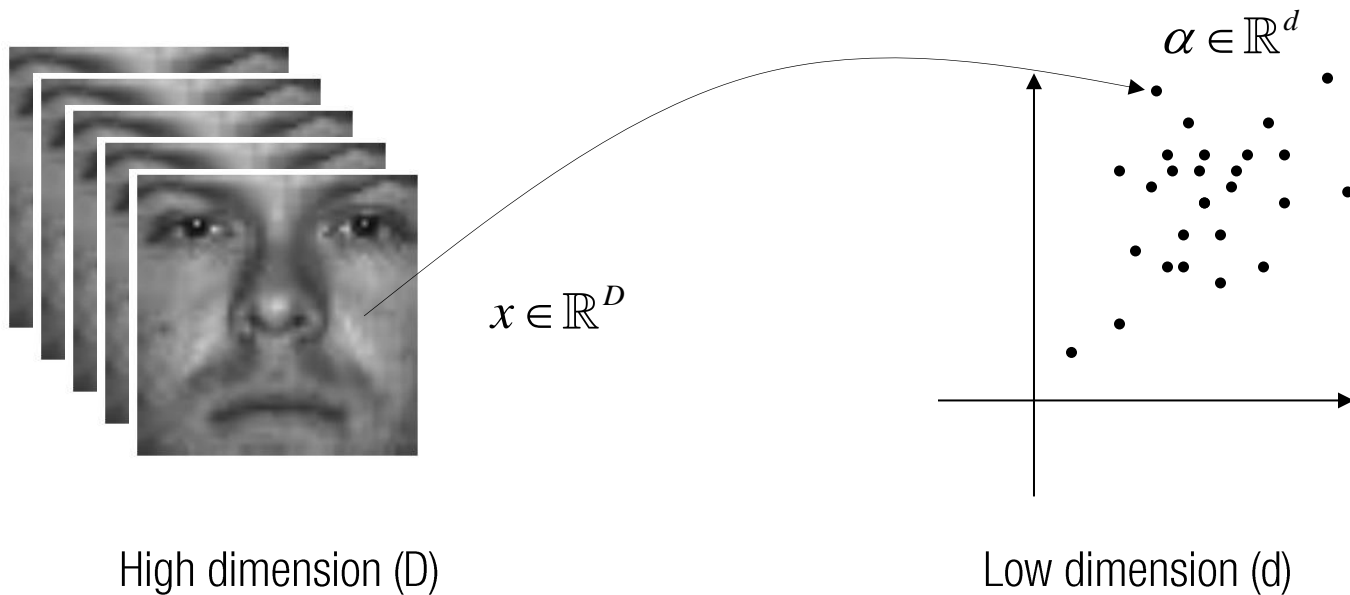
Observation: not all pixels are equally informative to detect a face



Average image

# *STRUCTURED APPEARANCE*

Idea: face images are highly correlated and can be represented in a low-dimensional subspace.



# LINEAR BASIS

Face = Mean face  $+ \alpha_1$  Basis 1  $+ \alpha_2$  Basis 2  $+ \alpha_3$  Basis 3  $+ \dots$

$y = m + \alpha_1 b_1 + \alpha_2 b_2 + \alpha_3 b_3 + \dots$

# LINEAR BASIS



Face

Mean face

Basis 1

Basis 2

Basis 3

$$y = m + \alpha_1 b_1 + \alpha_2 b_2 + \alpha_3 b_3 + \dots$$

Below the equation, vertical bars represent vectors. An orange bar is under  $y$ , a gray bar is under  $m$ , a dark blue bar is under  $\alpha_1$ , a medium blue bar is under  $\alpha_2$ , and a light blue bar is under  $\alpha_3$ .

# LINEAR BASIS



Face

Mean face

Basis 1

Basis 2

Basis 3

$$y \approx m + B\alpha$$

The diagram illustrates the linear basis equation. On the left, a vertical orange bar represents the input  $y$ . This is followed by an approximation symbol  $\approx$ . To the right of the symbol is a vertical gray bar representing the mean face  $m$ . A plus sign  $+$  follows. To the right of the plus sign is a matrix  $B$ , represented by a bracket containing three vertical bars of decreasing height (dark blue, medium blue, light blue) and an ellipsis  $\dots$ . To the right of matrix  $B$  is a vector  $\alpha$ , represented by a purple box containing  $\alpha_1$ , a vertical ellipsis  $\vdots$ , and  $\alpha_d$ . Above the matrix  $B$ , the terms  $+\alpha_1 b_1$ ,  $+\alpha_2 b_2$ , and  $+\alpha_3 b_3$  are shown, corresponding to the basis images in the top row. The entire equation is followed by  $+\dots$ .

# RECONSTRUCTION FROM LINEAR BASIS



Face

Mean face

Basis 1

Basis 2

Basis 3

$$y = m + \alpha_1 b_1 + \alpha_2 b_2 + \alpha_3 b_3 + \dots$$

$$y \approx m + \begin{bmatrix} | & | & | & \dots \\ \text{Basis} & & & \\ | & | & | & \dots \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{bmatrix}$$

$$y \approx m + B \alpha$$

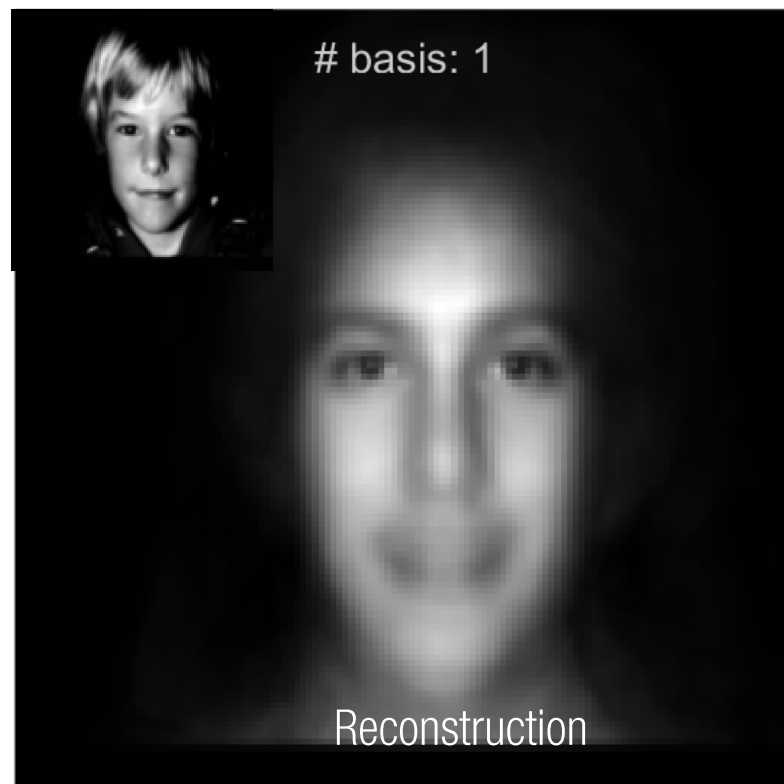
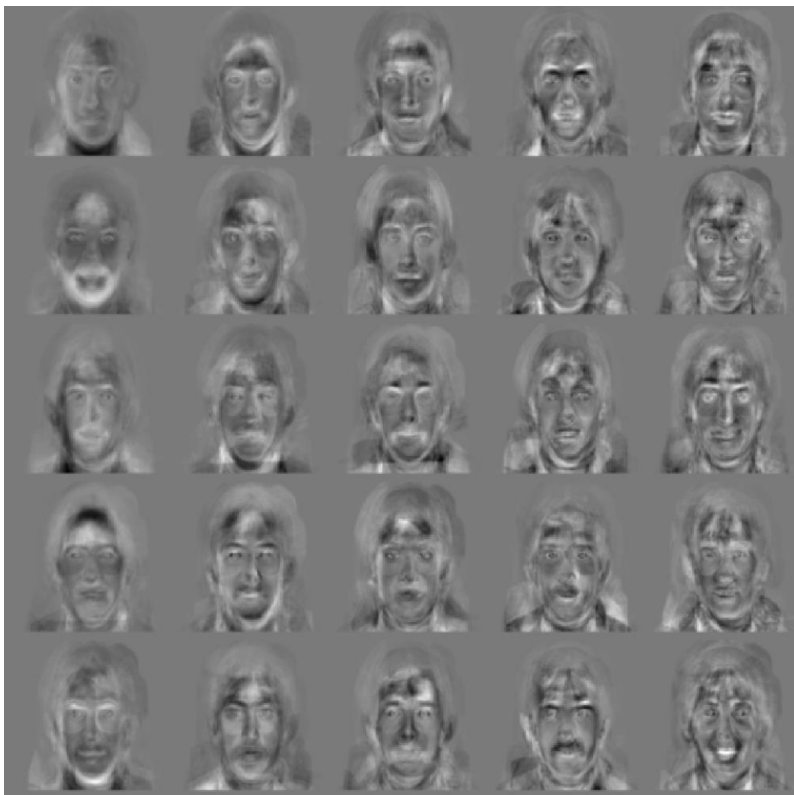
Mean
Basis
Coefficient

$$\alpha^* = \underset{\alpha}{\text{minimize}} \|y - m - B\alpha\|^2$$

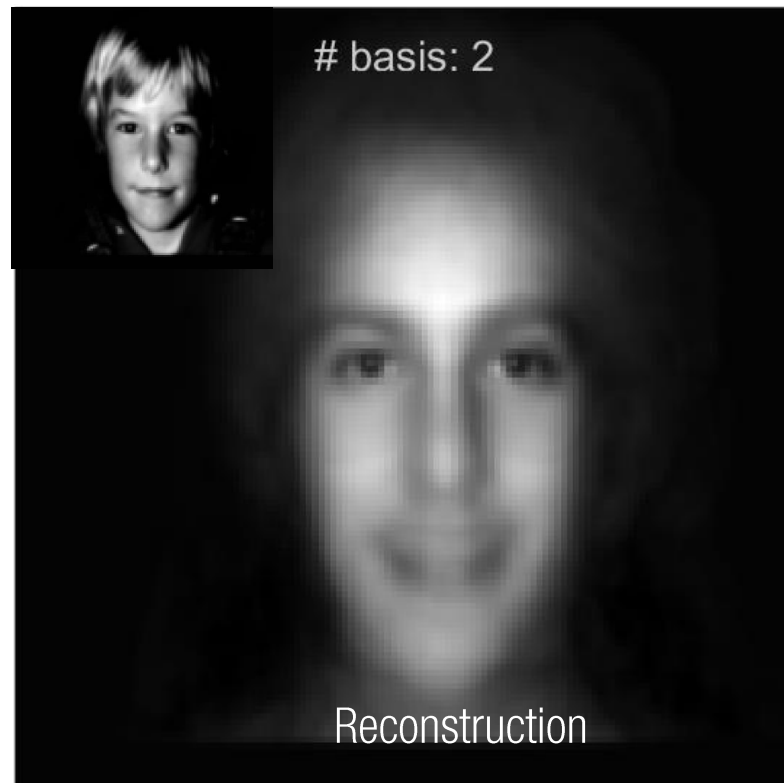
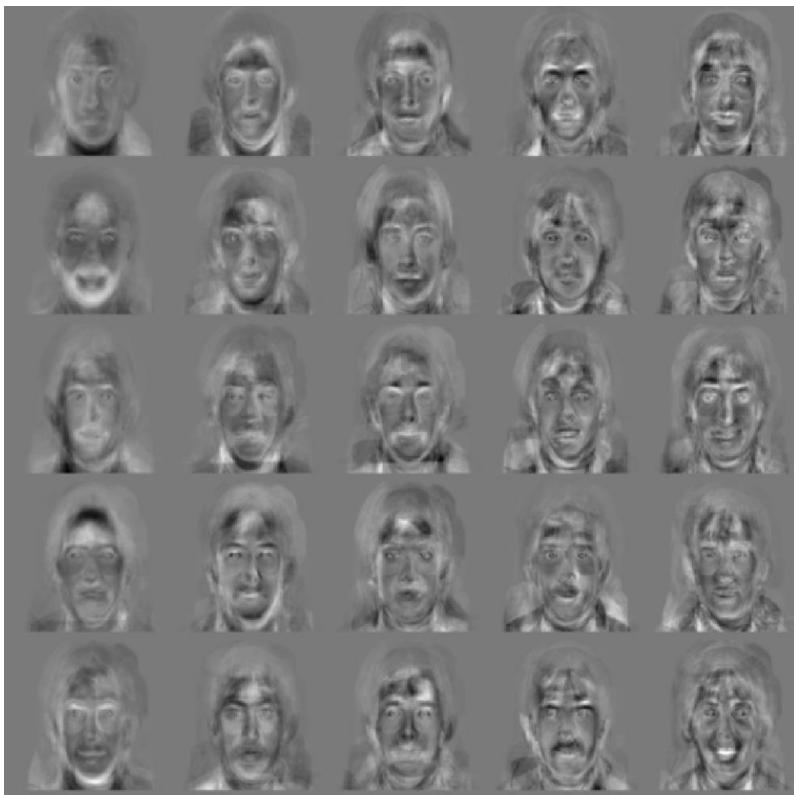
Cf)  $\text{minimize} \|x - y\|^2$   
 Template matching



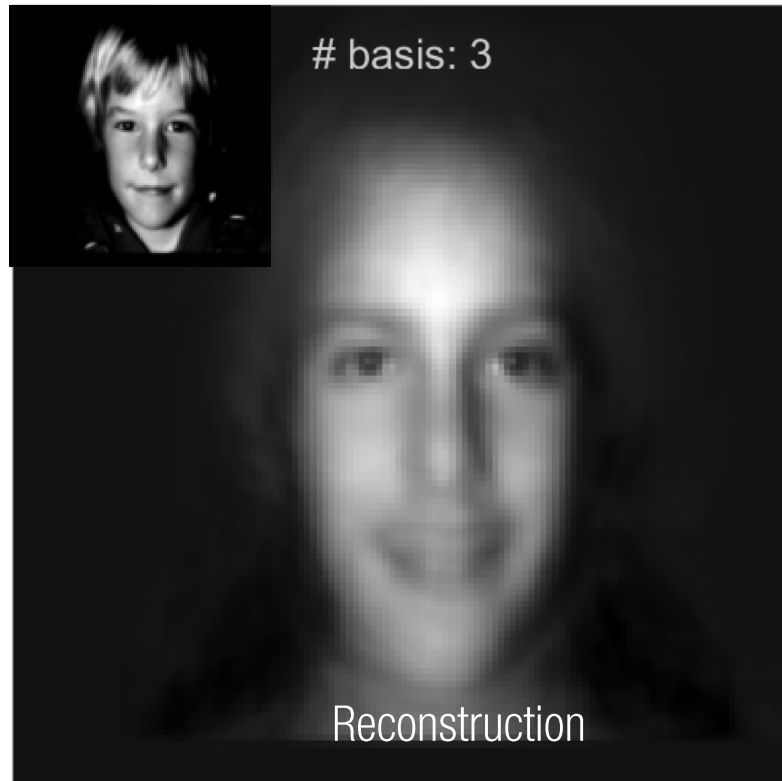
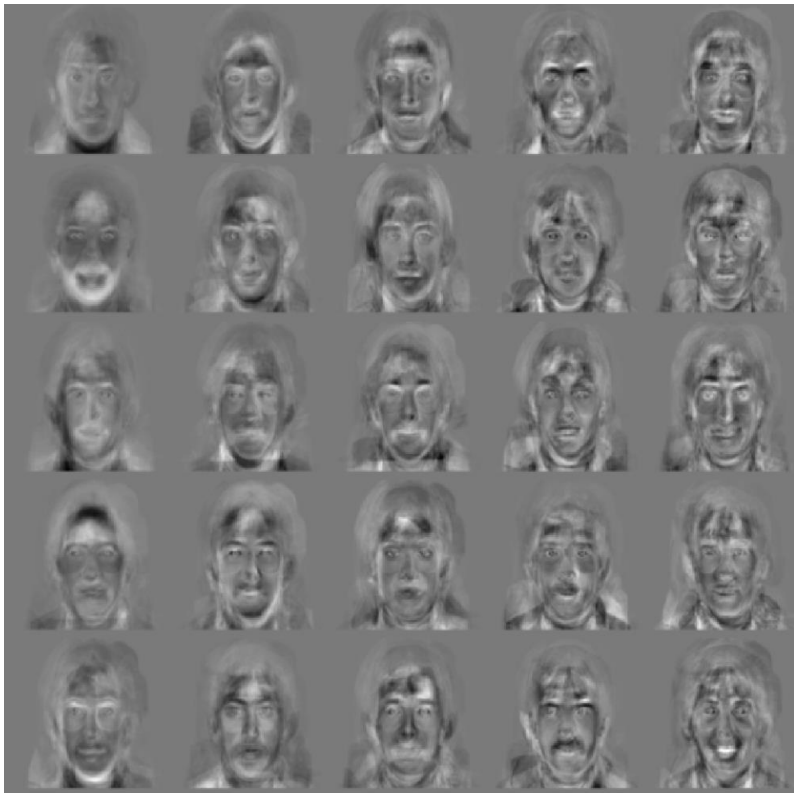
# *RECONSTRUCTION EXPRESSIBILITY*



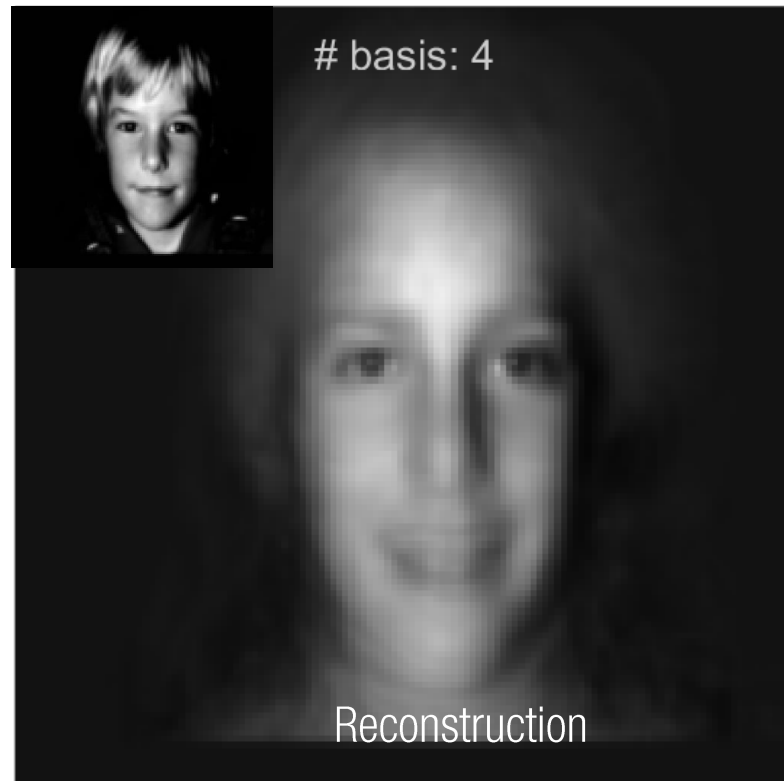
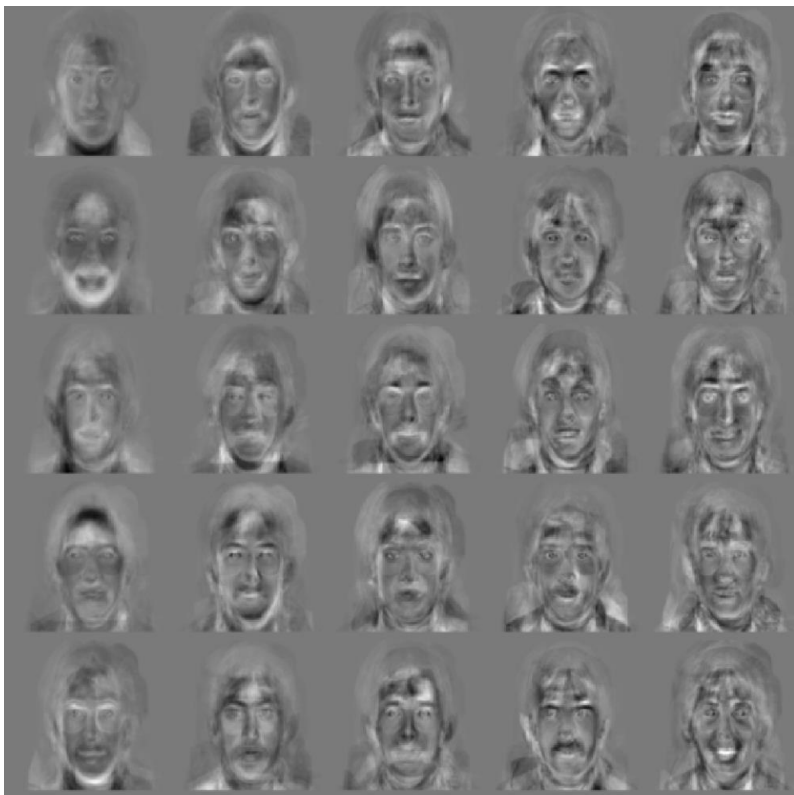
# *RECONSTRUCTION EXPRESSIBILITY*



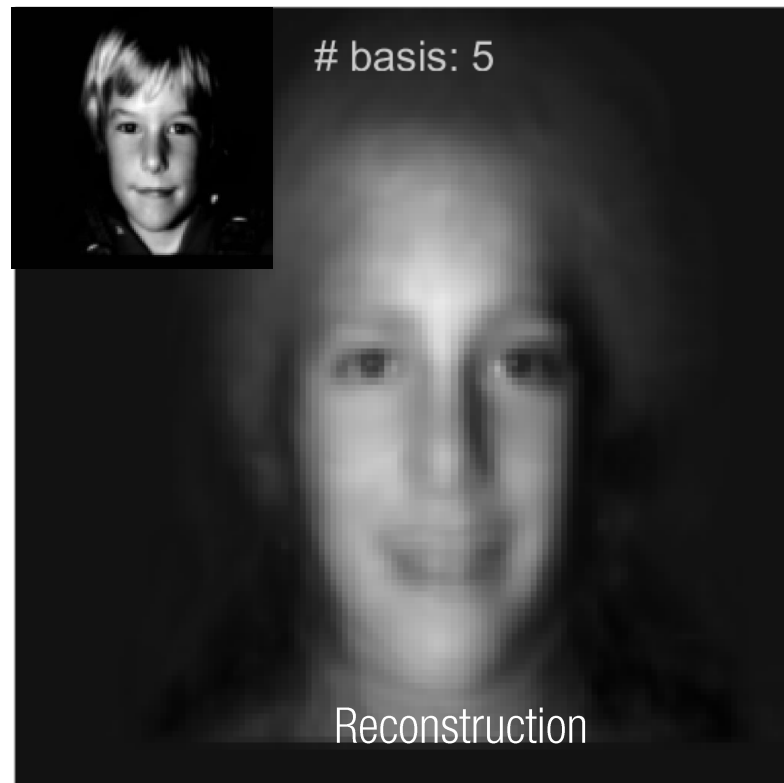
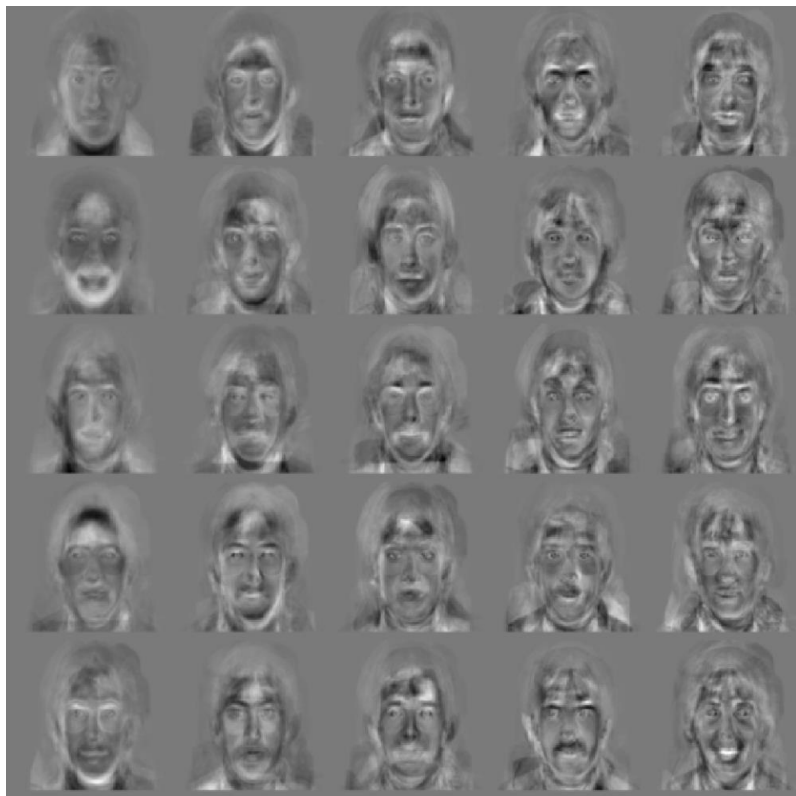
# *RECONSTRUCTION EXPRESSIBILITY*



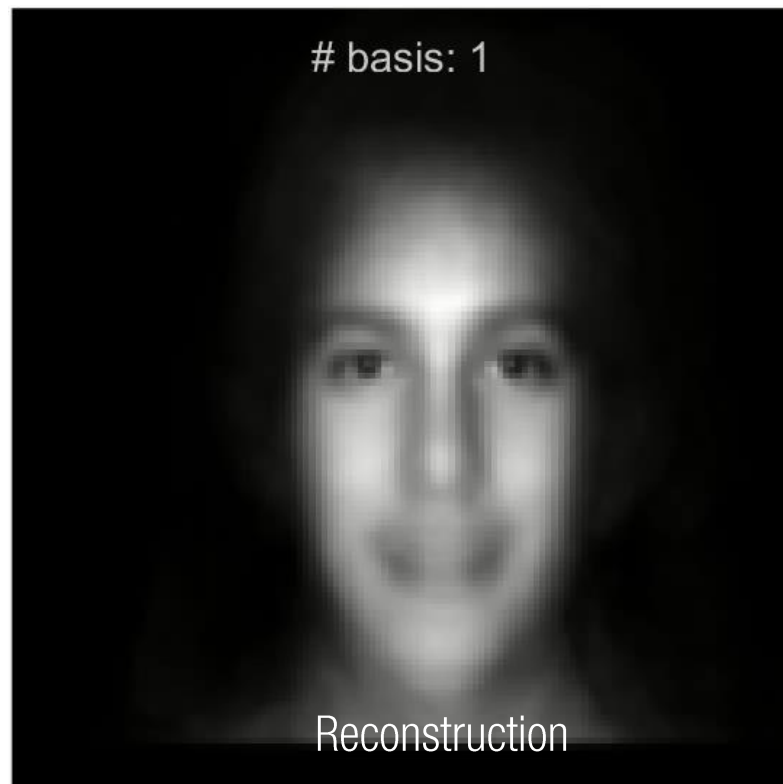
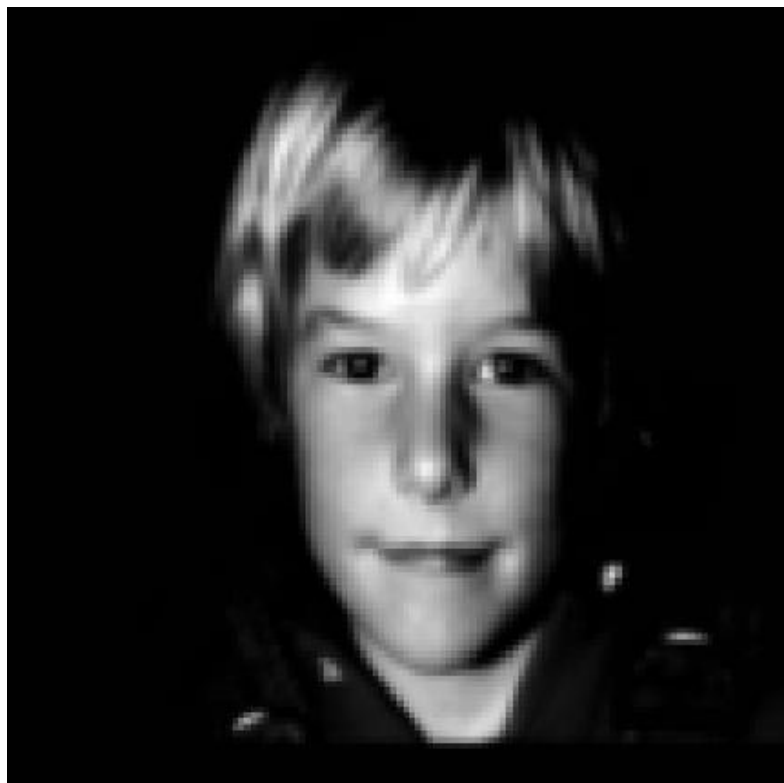
# *RECONSTRUCTION EXPRESSIBILITY*



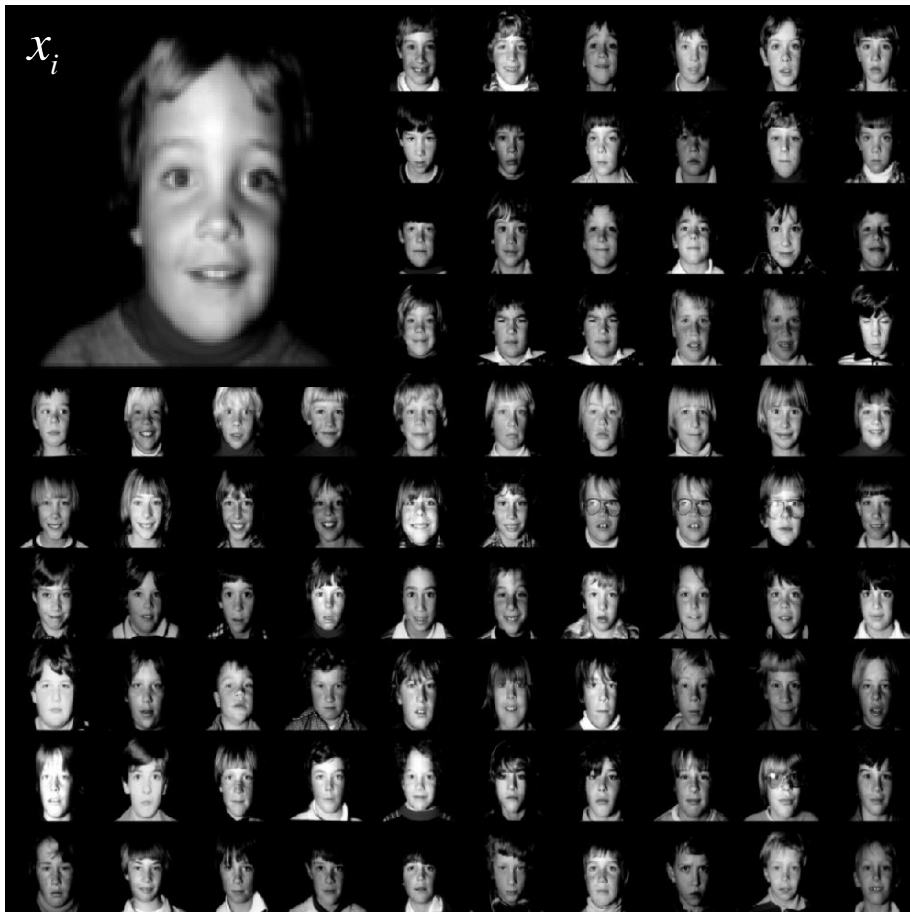
# *RECONSTRUCTION EXPRESSIBILITY*



# *RECONSTRUCTION EXPRESSIBILITY*

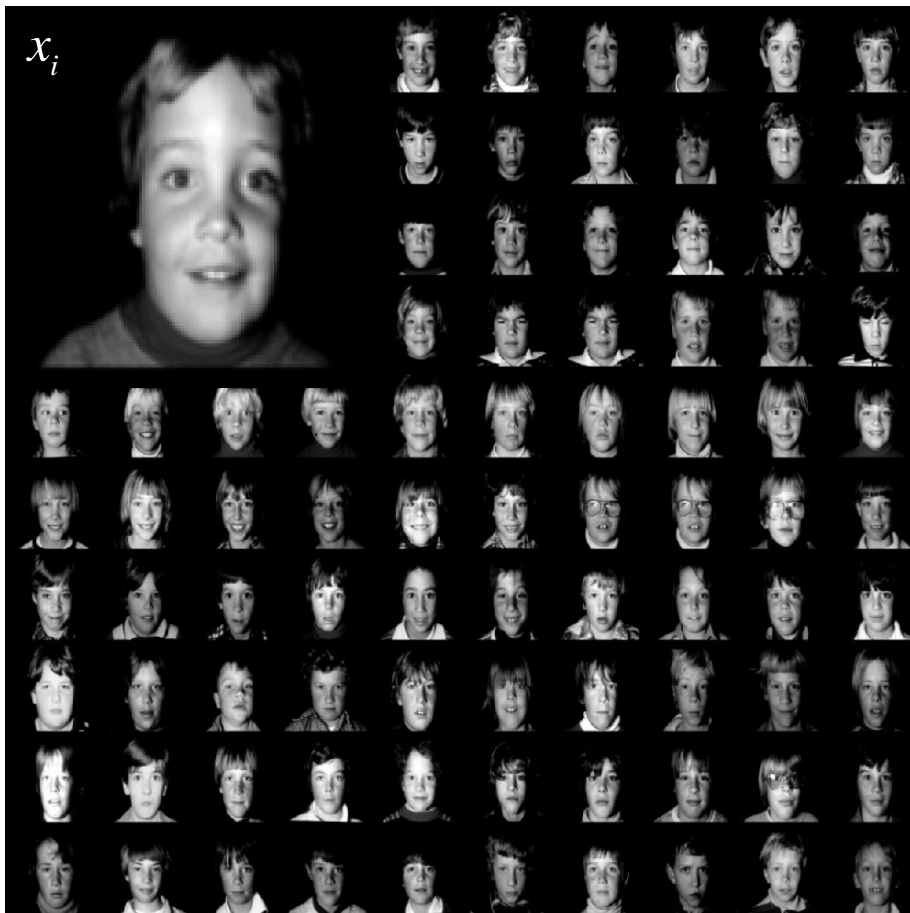


# HOW TO COMPUTE MEAN AND BASIS FROM DATABASE?



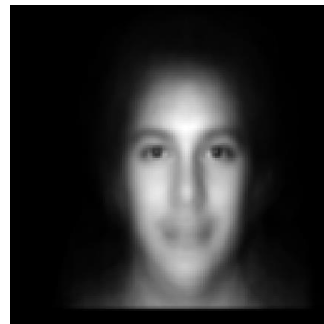
$$m^*, B^*, \alpha^* = \underset{m, B, \alpha}{\text{minimize}} \|y - m - B\alpha\|^2$$

# HOW TO COMPUTE MEAN?



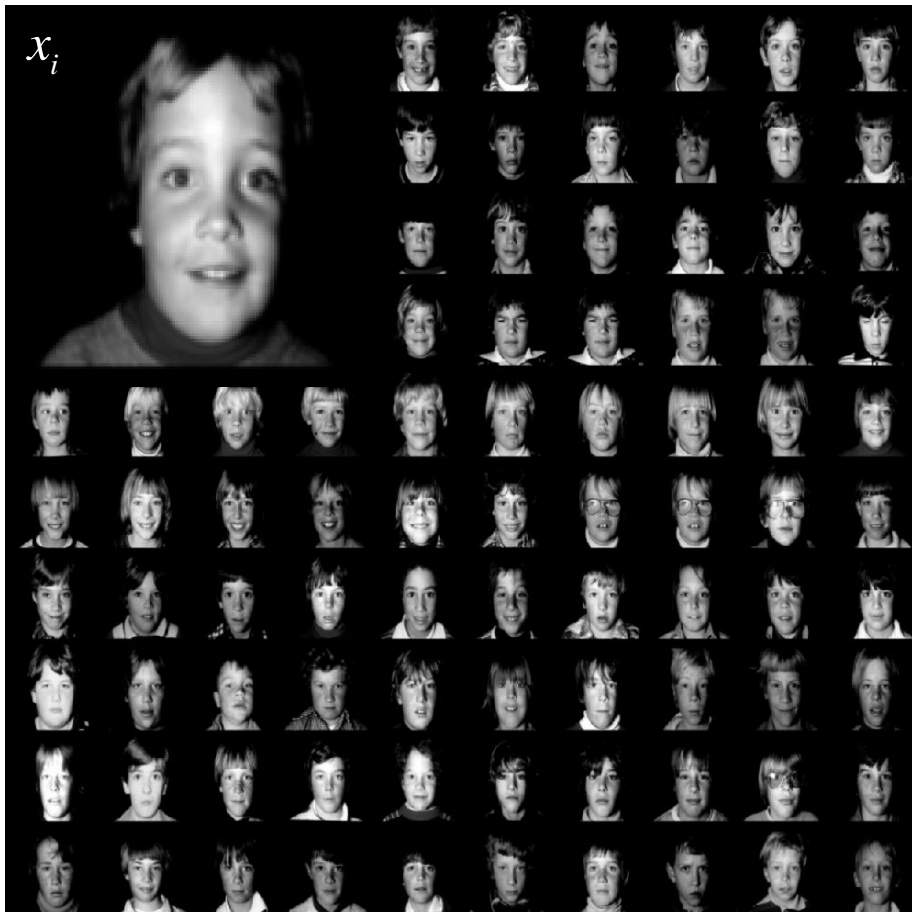
$$m^*, B^*, \alpha^* = \underset{m, B, \alpha}{\text{minimize}} \|y - m - B\alpha\|^2$$

$$m = \frac{1}{n} \sum_i^n x_i$$



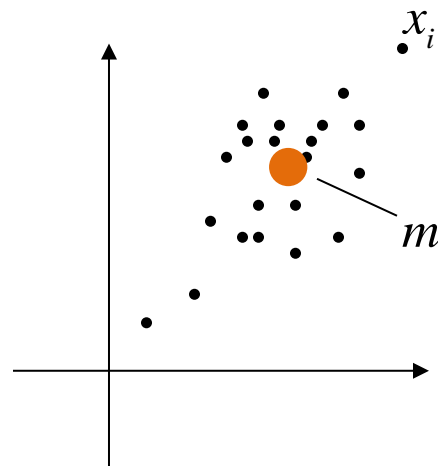
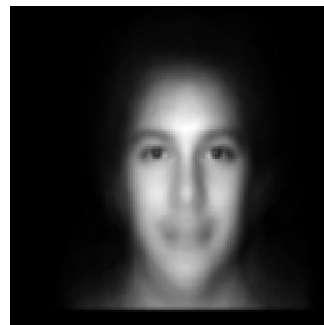


# HOW TO COMPUTE MEAN?



$$m^*, B^*, \alpha^* = \underset{m, B, \alpha}{\text{minimize}} \|y - m - B\alpha\|^2$$

$$m = \frac{1}{n} \sum_i^n x_i$$

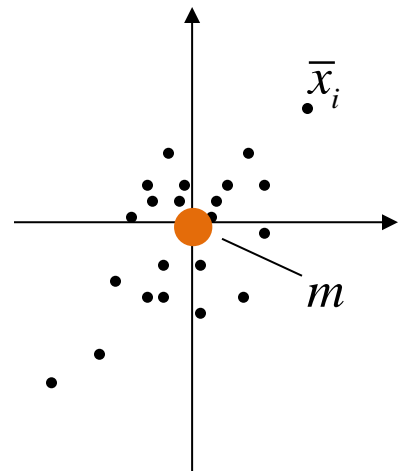


# MEAN SUBTRACTION



$$m^*, B^*, \alpha^* = \underset{m, B, \alpha}{\text{minimize}} \|y - m - B\alpha\|^2$$

$$m = \frac{1}{n} \sum_i^n x_i$$



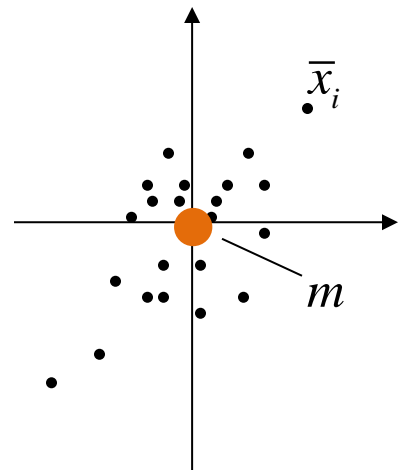
$$\bar{x}_i = x_i - m$$

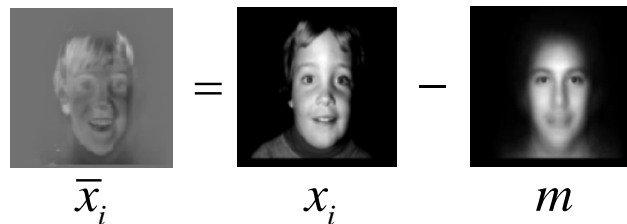
# HOW TO COMPUTER BASIS?



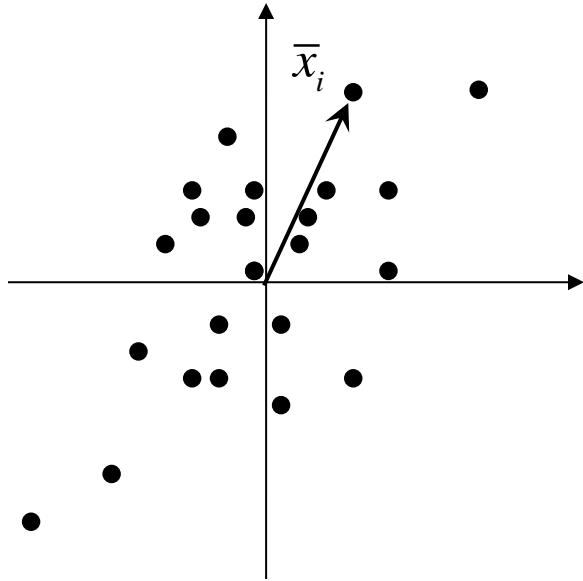
$$m^*, B^*, \alpha^* = \underset{m, B, \alpha}{\text{minimize}} \|y - m - B\alpha\|^2$$

$$m = \frac{1}{n} \sum_i^n x_i$$

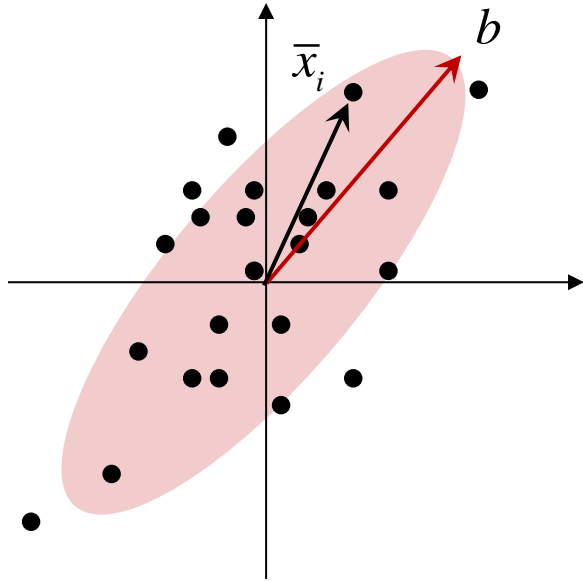



$$\bar{x}_i = x_i - m$$

# *PRINCIPAL AXIS*

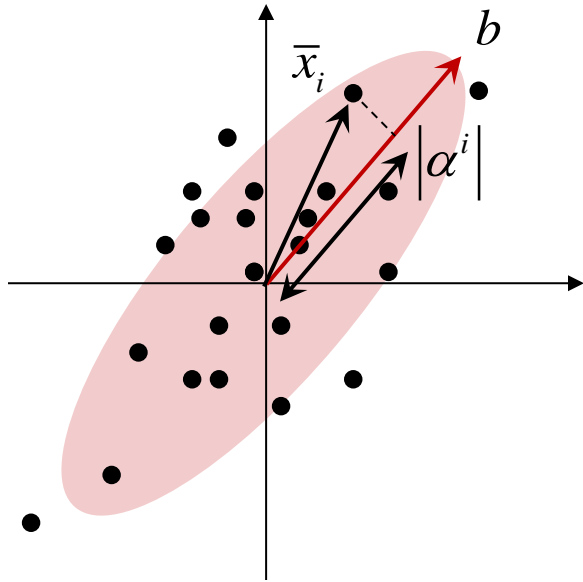


# *PRINCIPAL AXIS*



Basis is the axis that represents the maximum data covariance.

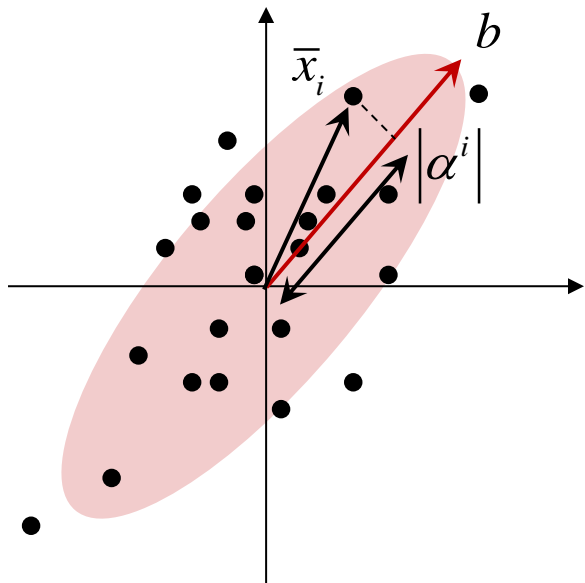
# PRINCIPAL AXIS



Basis is the axis that represents the maximum data covariance.

Coefficient  $\alpha^i = \frac{b \cdot \bar{x}_i}{\|b\|}$

# PRINCIPAL AXIS

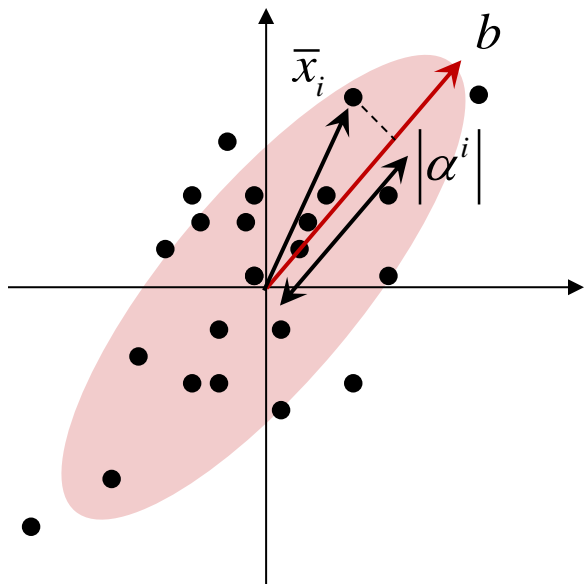


Basis is the axis that represents the maximum data covariance.

Coefficient  $\alpha^i = \frac{b \cdot \bar{x}_i}{\|b\|}$

$$b^* = \underset{b}{\text{maximize}} \sum_{i=1}^n (\alpha^i)^2$$

# PRINCIPAL AXIS



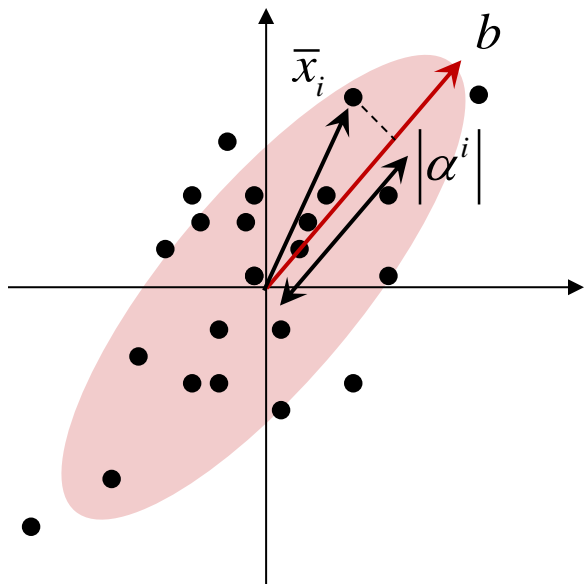
Basis is the axis that represents the maximum data covariance.

$$\text{Coefficient } \alpha^i = \frac{b \cdot \bar{x}_i}{\|b\|}$$

$$\begin{aligned} b^* &= \text{maximize}_b \sum_{i=1}^n (\alpha^i)^2 \\ &= \text{maximize}_b \sum_{i=1}^n \left( \frac{b \cdot \bar{x}_i}{\|b\|} \right)^2 \end{aligned}$$



# PRINCIPAL AXIS



Basis is the axis that represents the maximum data covariance.

$$\text{Coefficient } \alpha^i = \frac{b \cdot \bar{x}_i}{\|b\|}$$

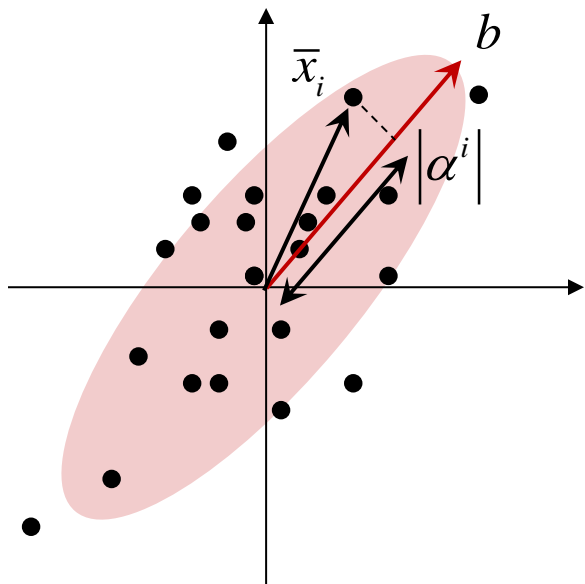
$$b^* = \underset{b}{\text{maximize}} \sum_{i=1}^n (\alpha^i)^2$$

$$= \underset{b}{\text{maximize}} \sum_{i=1}^n \left( \frac{b \cdot \bar{x}_i}{\|b\|} \right)^2$$

$$= \underset{b}{\text{maximize}} b^T \underbrace{X^T X}_{\text{Covariance matrix}} b$$

$$\text{where } x = \begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix}$$

# PRINCIPAL AXIS



Basis is the axis that represents the maximum data covariance.

$$\text{Coefficient } \alpha^i = \frac{b \cdot \bar{x}_i}{\|b\|}$$

$$b^* = \underset{b}{\text{maximize}} \sum_{i=1}^n (\alpha^i)^2$$

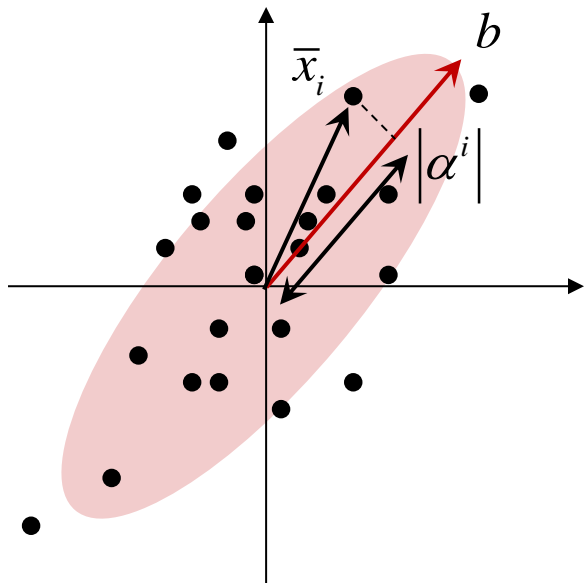
$$= \underset{b}{\text{maximize}} \sum_{i=1}^n \left( \frac{b \cdot \bar{x}_i}{\|b\|} \right)^2$$

$$= \underset{b}{\text{maximize}} b^T \underbrace{X^T X}_{\text{Covariance matrix}} b$$

Covariance matrix

Solution is the eigenvector corresponding to the largest eigenvalue:  $b^* = \lambda_{\max}(X^T X)$

# PRINCIPAL AXIS



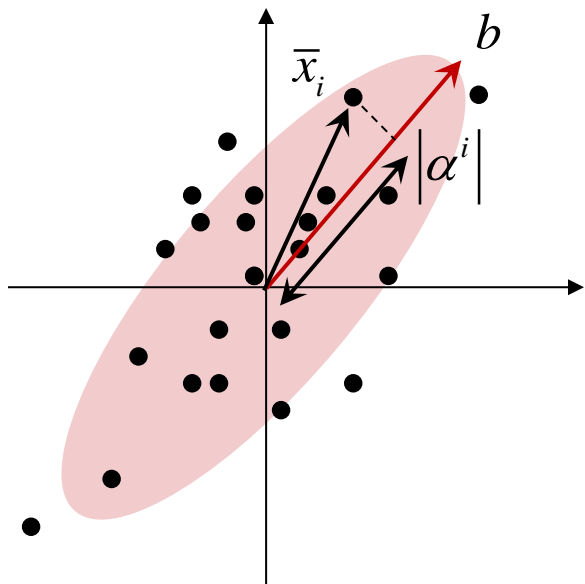
Basis is the axis that represents the maximum data covariance.

Coefficient  $\alpha^i = \frac{b \cdot \bar{x}_i}{\|b\|}$

$$\begin{bmatrix} \alpha^1 \\ \vdots \\ \alpha^n \end{bmatrix}_{n \times 1} = \begin{bmatrix} \bar{x}_1^T \\ \vdots \\ \bar{x}_n^T \end{bmatrix}_{n \times D} \begin{bmatrix} b \end{bmatrix}_{D \times 1}$$



# PRINCIPAL AXES



Basis is the axis that represents the maximum data covariance.

Coefficient  $\alpha_j^i = \frac{b_j \cdot \bar{x}_i}{\|b_j\|}$

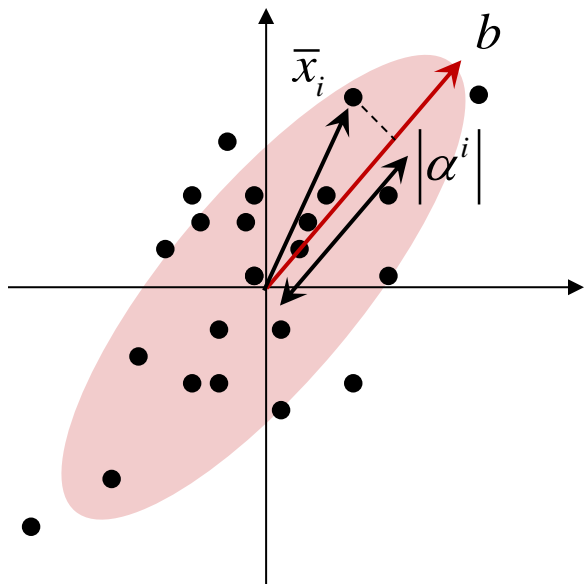
Orthogonal principal axes:  
first  $d$  largest eigenvectors

$$\begin{bmatrix} \alpha_1^1 & \dots & \alpha_d^1 \\ \vdots & & \vdots \\ \alpha_n^1 & \dots & \alpha_n^d \end{bmatrix}_{n \times d} = \begin{bmatrix} \bar{x}_1^T \\ \vdots \\ \bar{x}_n^T \end{bmatrix}_{n \times D} \begin{bmatrix} b_1 & \dots & b_d \end{bmatrix}_{D \times d}$$

$$d \ll D$$



# PCA: DIMENSIONAL REDUCTION

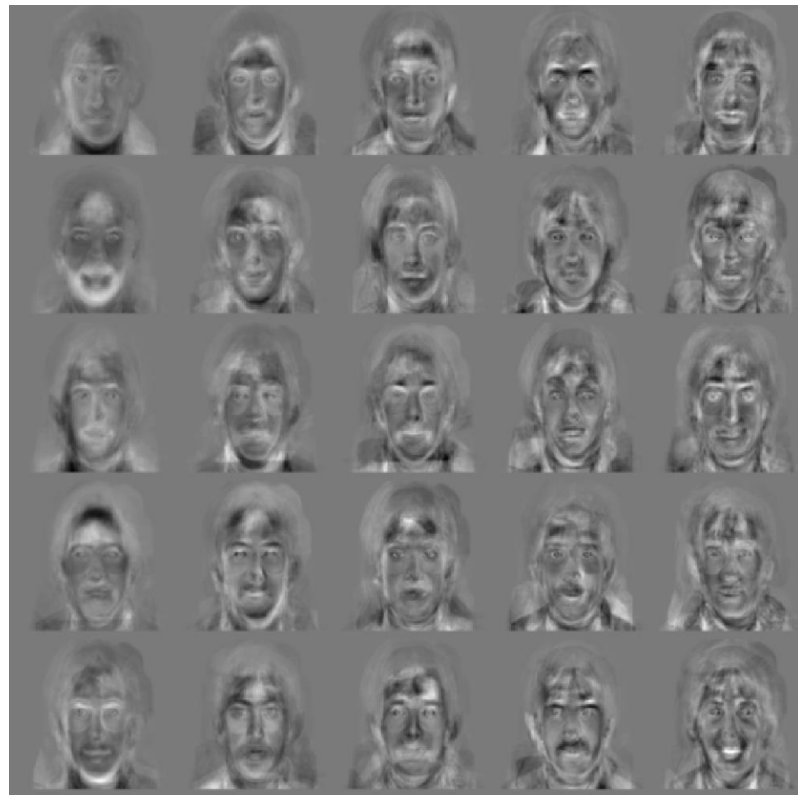


$$A = XB$$

$d \times n$                        $n \times D$                        $D \times d$

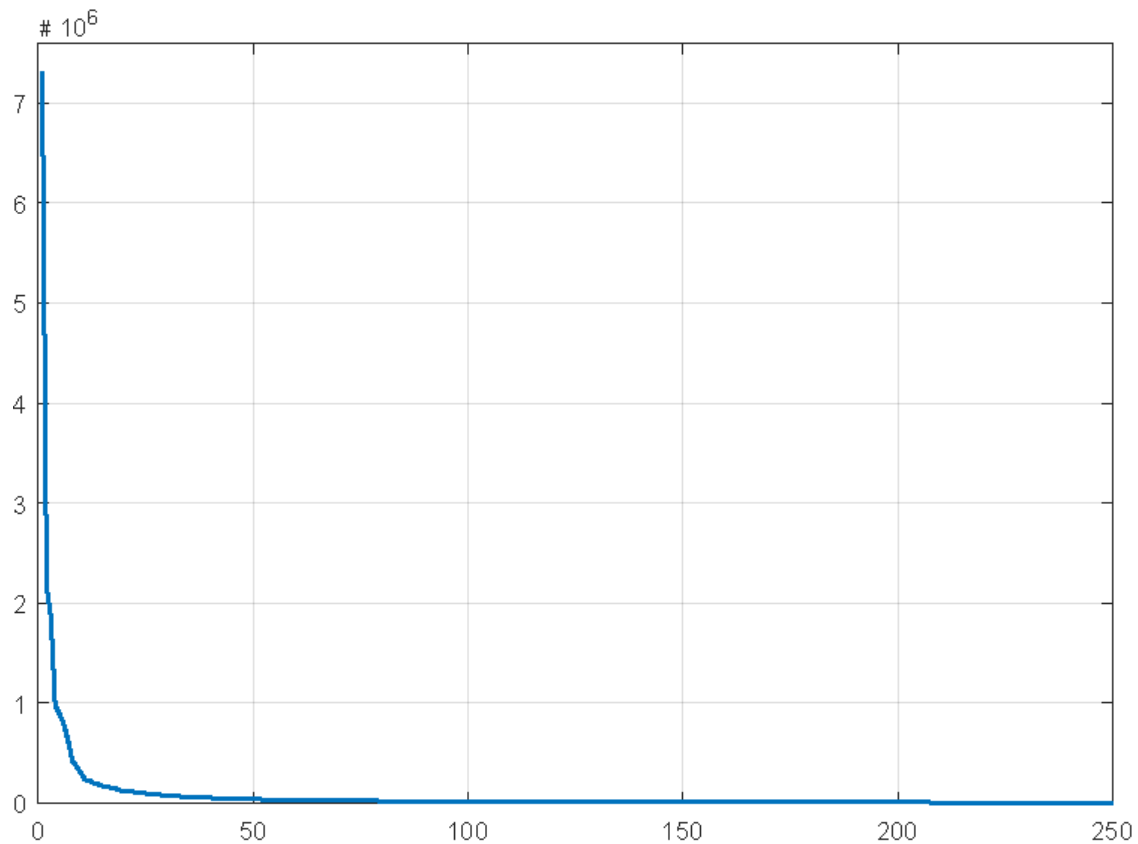
$d \ll D$

# *HOW TO COMPUTE BASIS?*

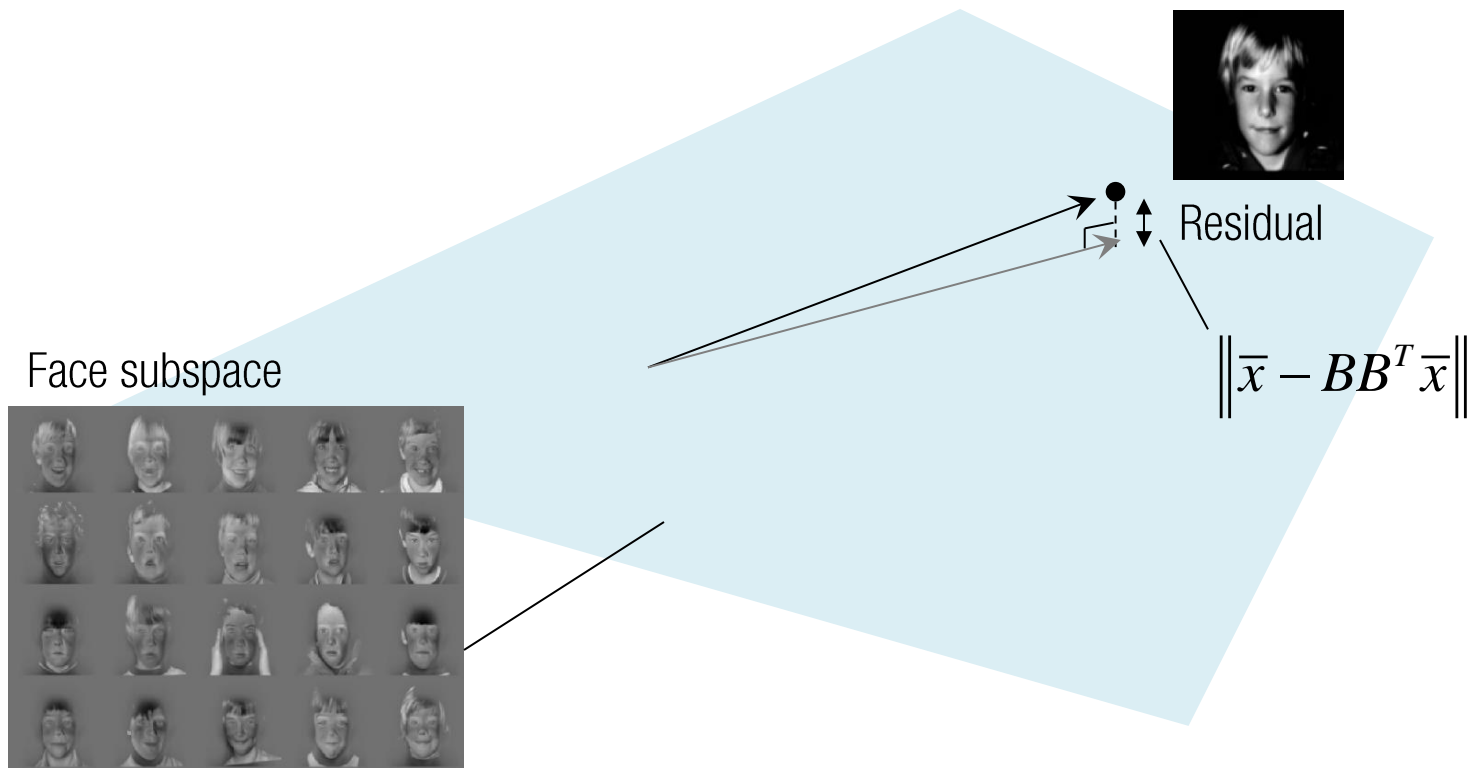


Set of basis vectors

# *HOW TO CHOOSE # OF BASIS VECTORS?*

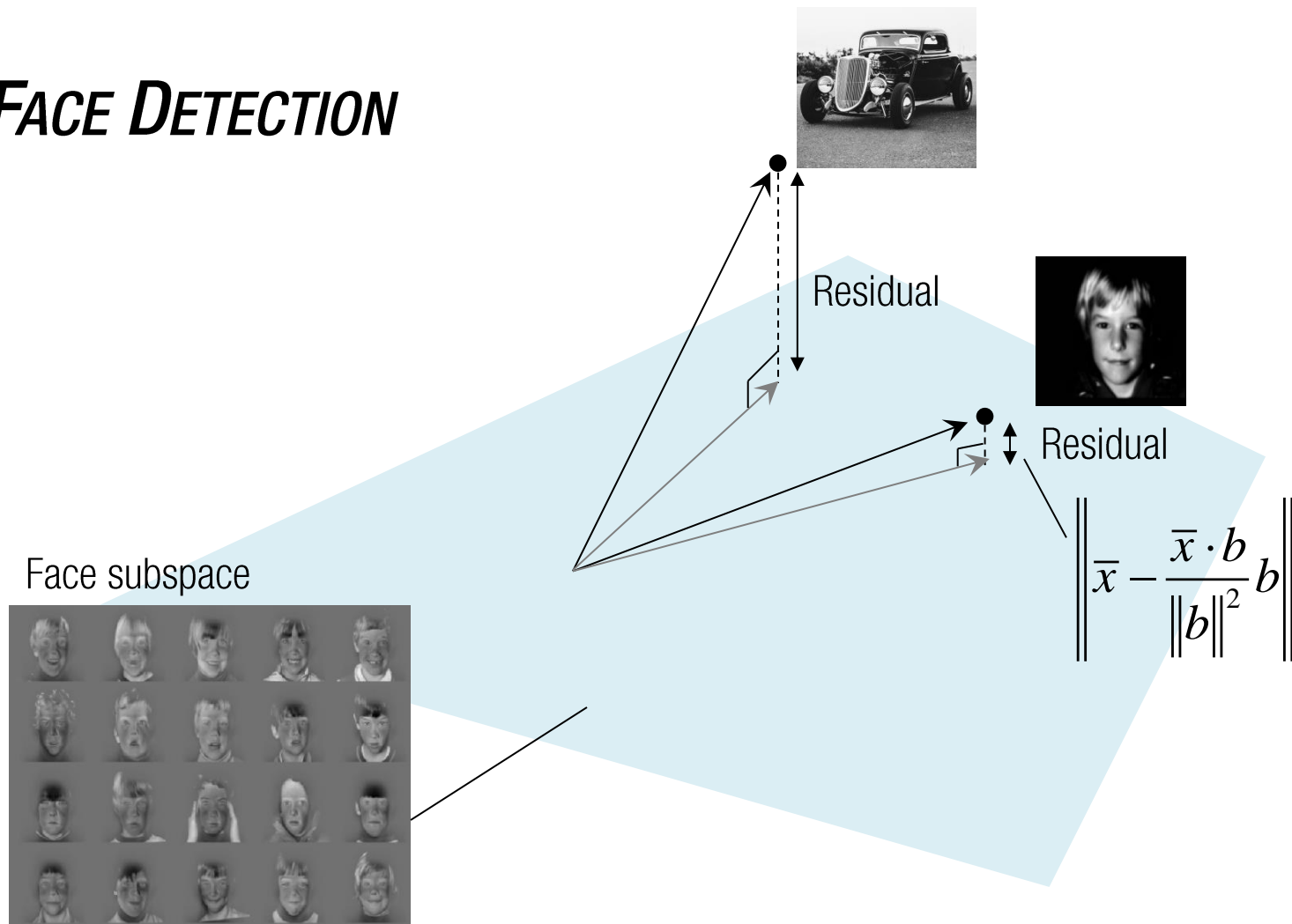


# FACE DETECTION

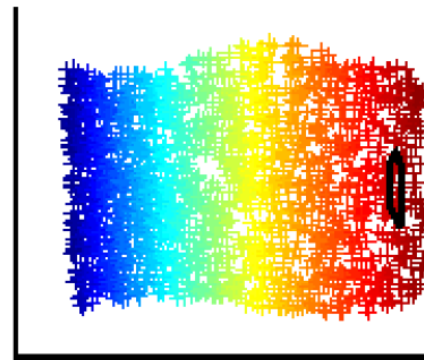
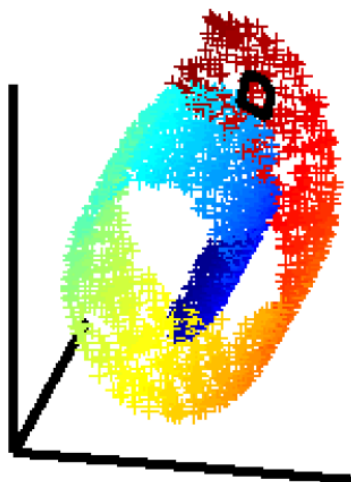
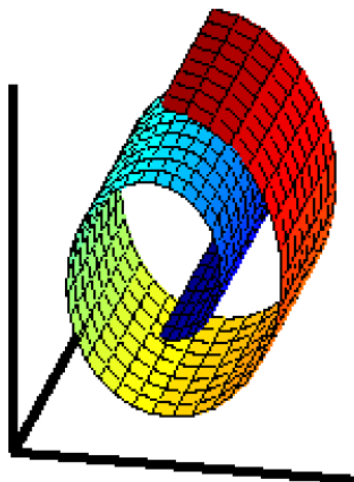




# FACE DETECTION



# *LIMITATION*



Object distribution does not follow Gaussian!

<https://www.youtube.com/watch?v=J0arU2PAMIs>