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)**

Template

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

High dimension (D)
e.g., D: 10,000 = 100 x 100
**SPACE OF APPEARANCE (FIXED SHAPE)**

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

Template

High dimension (D)
e.g., D: 10,000 = 100 x 100

Naïve face detection algorithm:

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.

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

$\alpha \in \mathbb{R}^d$
**LINEAR BASIS**

\[ y = m + \alpha_1 b_1 + \alpha_2 b_2 + \alpha_3 b_3 + \cdots \]
**Linear Basis**

\[
\begin{align*}
\text{Face} & = m + \alpha_1 b_1 + \alpha_2 b_2 + \alpha_3 b_3 + \cdots \\
y & = m + \alpha_1 b_1 + \alpha_2 b_2 + \alpha_3 b_3 + \cdots 
\end{align*}
\]
**Linear Basis**

\[ y = m \]

\[ y \approx m + B \alpha \]

\[ y = m + \alpha_1 b_1 + \alpha_2 b_2 + \alpha_3 b_3 + \ldots \]
**Reconstruction from Linear Basis**

\[ y \approx m + B \alpha \]

\[ y = m + \sum \alpha_i b_i \]

\[ \alpha^* = \min_{\alpha} \| y - m - B\alpha \|^2 \]

Cf) \minimize \| x - y \|^2

Template matching
Reconstruction Expressibility
RECONSTRUCTION EXPRESSIBILITY
Reconstruction Expressibility

# basis: 3

Reconstruction
Reconstruction Expressibility
Reconstruction Expressibility
RECONSTRUCTION EXPRESSIBILITY

# basis: 1

Reconstruction
**How to Compute Mean and Basis from Database?**

\[ m^*, B^*, \alpha^* = \text{minimize} \| y - m - B\alpha \|^2_{m,B,\alpha} \]
**How to Compute Mean?**

\[ m^*, B^*, \alpha^* = \min_{m, B, \alpha} \| y - m - B\alpha \|^2 \]

\[ m = \frac{1}{n} \sum_{i} x_i \]
How to Compute Mean?

$$m^*, B^*, \alpha^* = \min_{m, B, \alpha} \| y - m - B\alpha \|^2$$

$$m = \frac{1}{n} \sum_{i=1}^{n} x_i$$
Mean Subtraction

\[ m^*, B^*, \alpha^* = \min_{m, B, \alpha} \| y - m - B\alpha \|^2 \]

\[ m = \frac{1}{n} \sum_{i} x_i \]

\[ \bar{x}_i = x_i - m \]
How to Compute Basis?

\[ m^*, B^*, \alpha^* = \min_{m,B,\alpha} \| y - m - B\alpha \|^2 \]

\[ m = \frac{1}{n} \sum_{i} x_i \]

\[ \bar{x}_i = x_i - m \]
PRINCIPAL AXIS

\[ \bar{x}_i \]
Principal Axis

Basis is the axis that represents the maximum data covariance.
**Principal Axis**

Basis is the axis that represents the maximum data covariance.

\[
\alpha^i = \frac{b \cdot \bar{x}_i}{\|b\|}
\]
**Principal Axis**

Basis is the axis that represents the maximum data covariance.

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

\[ b^* = \max_b \sum_{i=1}^{n} (\alpha^i)^2 \]
Basis is the axis that represents the maximum data covariance.

Coefficient

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

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

$$= \max_b \sum_{i=1}^{n} \left( \frac{b \cdot \bar{x}_i}{\|b\|} \right)^2$$
**Principal Axis**

Basis is the axis that represents the maximum data covariance.

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

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

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

$= \max_b b^T X^T Xb$

Covariance matrix

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.

Coefficient \( \alpha^i = \frac{b \cdot \overline{x}_i}{\|b\|} \)

\( b^* = \max_b \sum_{i=1}^{n} (\alpha^i)^2 \)

\( = \max_b \sum_{i=1}^{n} \left( \frac{b \cdot \overline{x}_i}{\|b\|} \right)^2 \)

\( = \max_b b^T X^T X b \)

Covariance matrix

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

Basis is the axis that represents the maximum data covariance.

Coefficient \[ \alpha^i = \frac{\mathbf{b} \cdot \overline{x}_i}{\|\mathbf{b}\|} \]

\[
\begin{bmatrix}
\alpha^1 \\
\vdots \\
\alpha^n
\end{bmatrix} =
\begin{bmatrix}
\overline{x}_1^T \\
\vdots \\
\overline{x}_n^T
\end{bmatrix}
\begin{bmatrix}
b \\
\vdots \\
b
\end{bmatrix}
\]

nx1 \times nxD \times Dx1
**Principal Axes**

Basis is the axis that represents the maximum data covariance.

Coefficient\[ \alpha^i_j = \frac{b_j \cdot \bar{x}_i}{\|b_j\|} \]

Orthogonal principal axes: first \(d\) largest eigenvectors

\[
\begin{bmatrix}
\alpha^1_1 & \alpha^d_1 \\
\vdots & \vdots \\
\alpha^1_n & \alpha^d_n
\end{bmatrix}
\begin{bmatrix}
\bar{x}_1^T \\
\vdots \\
\bar{x}_n^T
\end{bmatrix}
= 
\begin{bmatrix}
b_1 \\
\vdots \\
b_d
\end{bmatrix}
\]

\(d \ll D\)
**PCA: Dimensional Reduction**

\[ \mathbf{A} = \mathbf{X} \mathbf{B} \]

- \( \mathbf{A} \) is of size \( d \times n \)
- \( \mathbf{X} \) is of size \( n \times D \)
- \( \mathbf{B} \) is of size \( D \times d \)

\[ d \ll D \]
How to Compute Basis?

Set of basis vectors
How to choose # of Basis Vectors?
**Face Detection**

\[
\| \bar{x} - BB^T \bar{x} \|
\]

Face subspace

Residual
FACE DETECTION

Residual

Face subspace

\[ \left\| \overline{x} - \frac{\overline{x} \cdot b}{\|b\|^2} b \right\| \]
LIMITATION

Object distribution does not follow Guassian!
https://www.youtube.com/watch?v=J0arU2PAMIs