Lecture 6 : Image & Recognition (2)

- Lecture 6 : Image & Recognition (2)
    - Motivations
    - Methods
        - PCA: Principal Component Analysis
    - - PCA Subspace Construction: **Eigenvectors** of the covariance matrix.
        - PCA Applications
        - Eigenfaces Problems
        - PCA Summary
        - LDA: Linear Discrimination Analysis
            - LDA Advantages
        - 2DPCA: Two-Dimensional Principal Component Analysis
            - 2DPCA Method
            - Disadvantage of 2DPCA
            - 2DPCA + PCA
            - 2DPCA Summary

Motivations

1. The large data size.
2. Some valuable information exists in the **latent sub-space**.

Methods

The key idea is to **combine features** using a **linear transformation**.

Linear combinations: simple to compute and analytically tractive.

$$y = U^T x \in \mathbb{R}^K, k << N$$

Detail:

1. PCA: Principal Component Analysis

PCA: Principal Component Analysis

- K-L transformation.
- Reduce the number of dimensions of a data set.
- Goal:
  - a. Feature extraction
  - b. Visualization
- Objective: Find basic vectors that
  - **Max** the **variance** retained projected data.
  - Give **uncorrelated** projected distributions.
  - **Min** the **least square reconstruction error**.
- **Similar** images are **near** each other.
- **Different** images are **far** from each other.
- PCA Subspace Construction: **Eigenvectors** of the covariance matrix.

PCA Applications

- A 256*256 pixel image => 65,536-dimensional image space.
- Reduce the dimensions.
- Find the nearest 'known' vector.

Example:

1. Training Set X and Test Set Y.
2. Total mean for every X is m_i, and $m = \frac{1}{|X|} \sum m$.
3. Total scatter matrix (Covariance Matrix) $S_t = E[(X - m)(X - m)^T] = \sum p_i (\sum p_j (X_j^i - m)(X_j^i - m)^T)$
4. Compute the eigenvalue and eigenvectors of S_t. $\lambda_i \phi_i = S_t \phi_i$.
5. Select the most principal components or eigenvectors. $r_\lambda = \frac{\sum_i^6 \lambda_i}{\sum_i^6 44\lambda_i} > 90\%$. r_λ is the ratio of eigenvalue sum of the selected comopnents to the total sum.
6. The projection transform W is $W = (\phi_1, \phi_2, \phi_3, ..., \phi_6)$. $X' = (X - m)W$, $Y' = (Y - m)W$. Result: 6 dimensions.

Eigenfaces Problems

1. Lighting.
2. Orientation.
3. Size.

PCA Summary

- Advantages
  - Reduce data dimensionality.
  - Satisfy the Minimal MSE Rule.
  - Eliminate the correlation of original data.
- Disadvantages
  - No distinguishing between shape and appearance.
  - No use class information.
  - Required transformation from image into a vector first.

LDA: Linear Discrimination Analysis

1. For X, compute within-class scatter matrix S_w and between-class scatter matrix S_b. $S_w = \sum p_i E[(X - m_i)(X - m_i)^T]$. $S_b = \sum p_i E[(m_i - m)(m_i - m)^T]$. p_i: prior probability of the i-th class. m_i: mean value of the i-th class.
2. Compute the eigenvectors of $S_w^{-1} S_b$.

Maximize the **component axes** for class-separation.

Good for **classification**.

LDA Advantages

Fisher criterion: not simple but very efficient **discriminative feature extraction** for the classification.
Further reduce the dimsion: rank of $S_w^{-1}S_b = c - 1$.

2DPCA: Two-Dimensional Principal Component Analysis

- No need to transform into a vector.
- Covariance matrix is constructed directly using original matrices.
- Smaller size of covariance matrix.
- $X \in \mathbb{R}^{M \times N}$, a unitary volumn vector $w \in \mathbb{R}^{N \times 1}$. Project X to w: $y = Xw$.
- Remaining question: how to find a good projection vector w.

2DPCA Method

1. Training and testing sets.
2. Express X by $X_j^i, \quad i = 1, 2;\ j = 1, ..., 5$ (2 classes, 5 samples for each class).
3. Compute the total scatter matrix $S_t = E[(X - M)^T(X - M)] = \sum p_i(\sum p_j^i(X_j^i - M)^T(X_j^i - M))$. X and M are $M \times N$ matrices. S_t is an $N \times N$ matrix.
4. Eigenvalue and eigenvectors of S_t, $\lambda_i \phi_i = S_t \phi_i$.
5. Select the most principal eigenvectors.
6. $W = (\phi_1, \phi_2, ..., \phi_d)$. $X' = (X - M)W \quad Y' = (Y - M)W$.
7. $X' = (X - M)V \quad V = (\phi_1, \phi_2, ..., \phi_N) \quad X = M + X'V^T$

Disadvantage of 2DPCA

The feature dimension of 2DPCA is much higher than PCA.

2DPCA + PCA

1. 2DPCA: $M \times N$ => $M \times d$.
2. PCA: $M \times d$ => $d' \times 1$.

2DPCA Summary

- Simpler and more straightforward.
- Better in **recognition accuracy**.
- More **efficient** than PCA.
- More **suitablel** for small sample size problems.
- **Covariance matrix** more **accurately**.