Tutorial 4

Qu Xiaofeng, Teaching Asistant
quxiaofeng.at.polyu@gmail.com

COMP435p
Biometrics Authentication

February 13, 2012

THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學

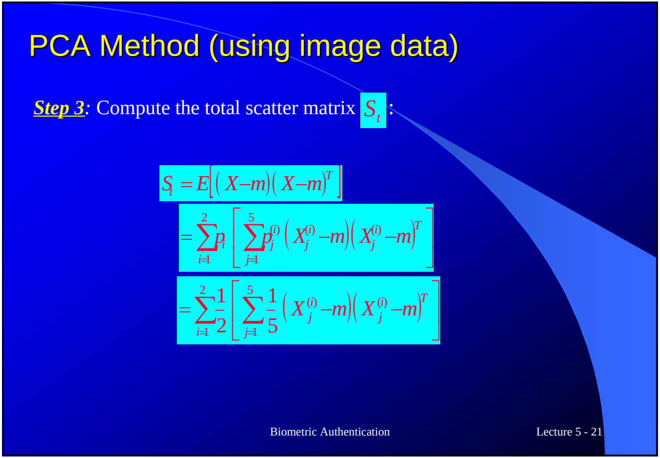
Table of Contents

Outline

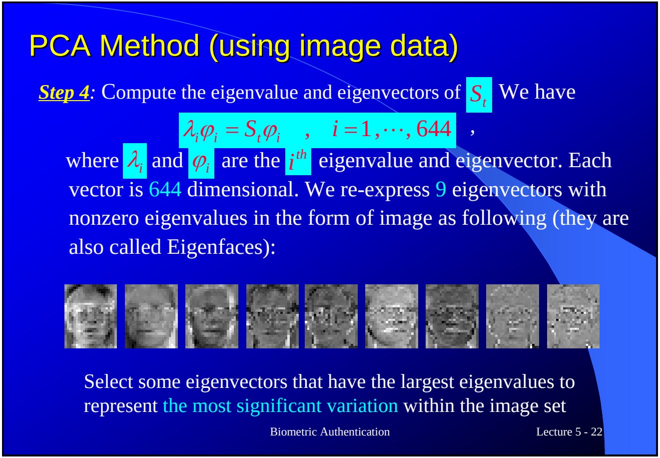
Problem 1.1 scatter matrix, eigenvector and eigenvalue

What is the scatter matrix (P5:21)? Understand what about eigenvector and eigenvalue as well as their functions? (P5:22)

Problem 1.1 scatter matrix, eigenvector and eigenvalue



Problem 1.1 scatter matrix, eigenvector and eigenvalue



Problem 1.1 scatter matrix, eigenvector and eigenvalue

"Detailed links(click)"

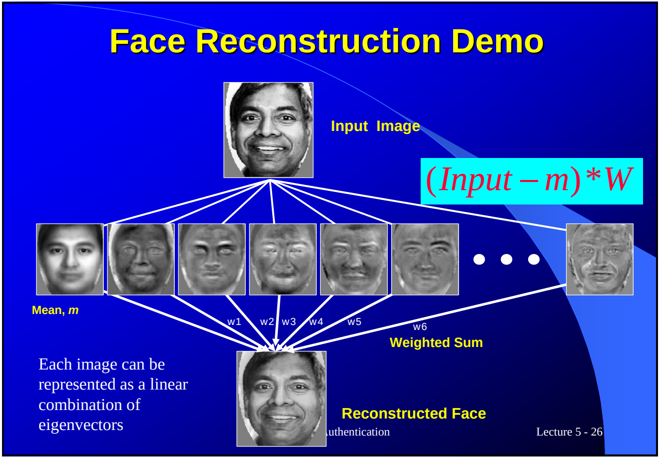
"PCA books(click)"

kolho3.tiera.ru

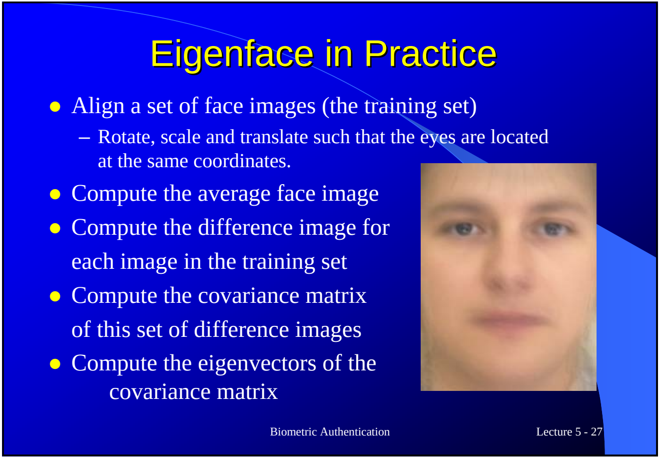
Problem 1.2 Eigenface

Understand the PCA application to facial recognition: Eigenface. (P5:26-30)

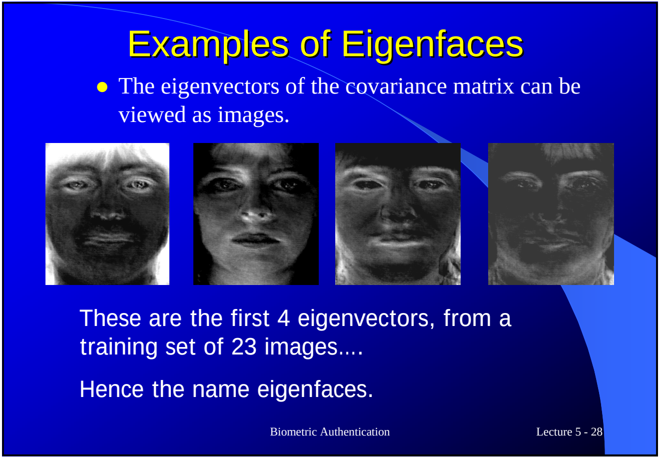
Problem 1.2 Eigenface



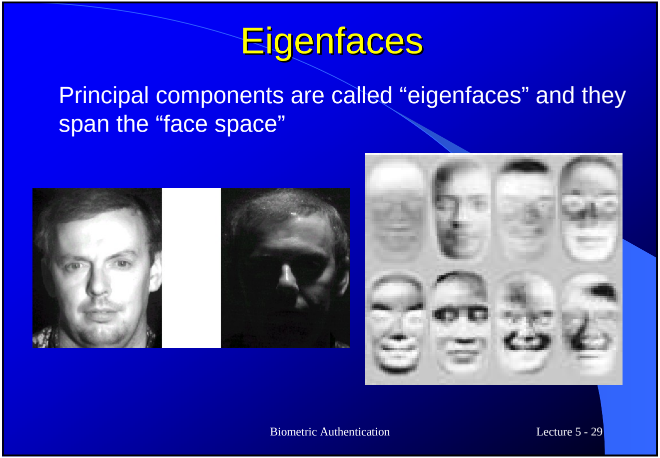
Problem 1.2 Eigenface



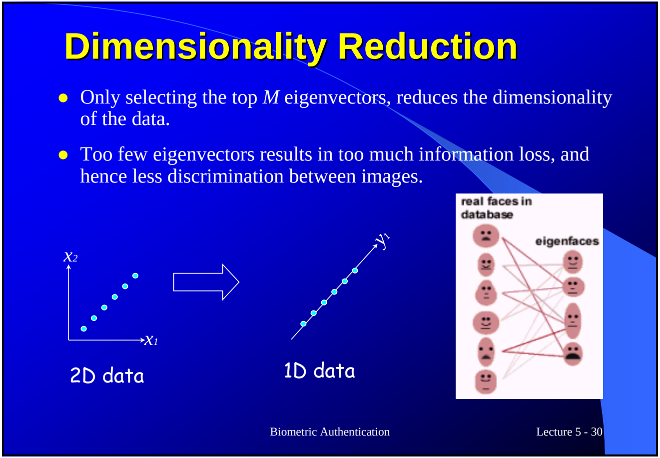
Problem 1.2 Eigenface



Problem 1.2 Eigenface



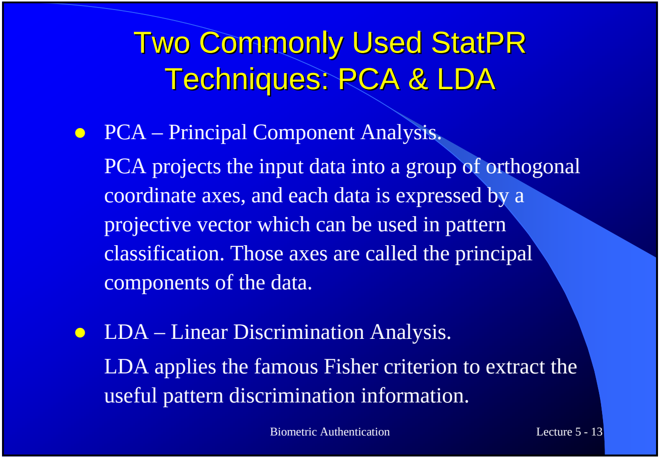
Problem 1.2 Eigenface



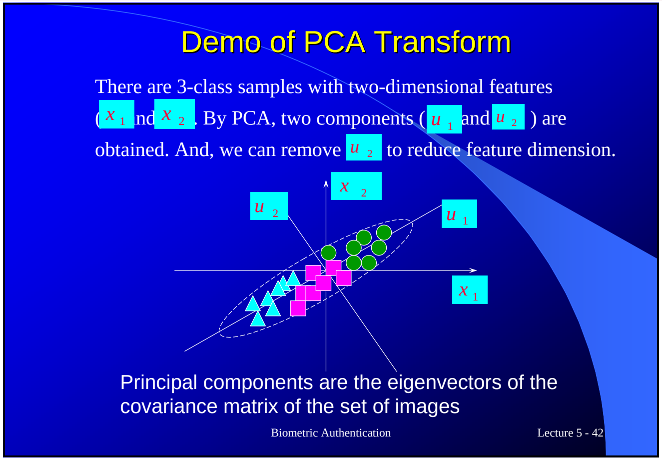
Problem 1.3 PCA and LDA

Compare two StatPR techniques, PCA and LDA (P5:13) and point out their main difference (P5:42-45)

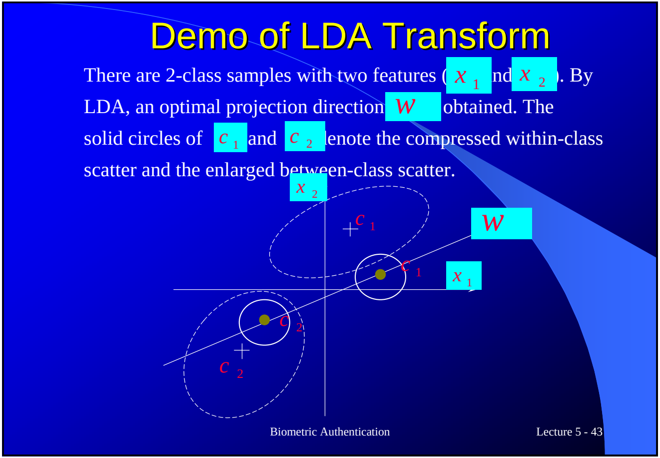
Problem 1.3 PCA and LDA



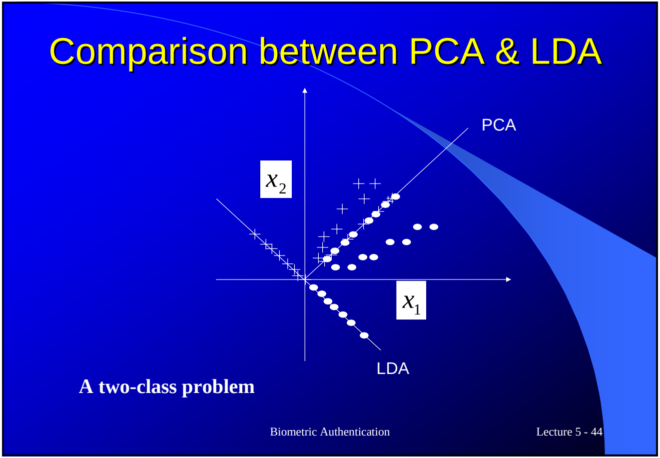
Problem 1.3 PCA and LDA



Problem 1.3 PCA and LDA



Problem 1.3 PCA and LDA



Problem 1.3 PCA and LDA

Eigenface & Fisherface

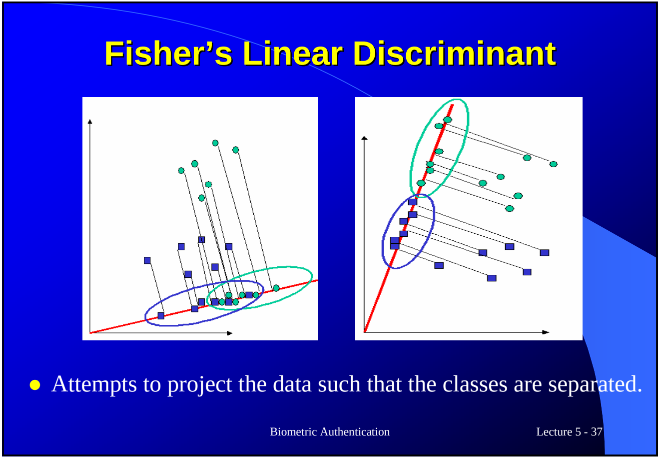
- *Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*
  - P. Belhumeur, J. Hespanha, D. Kriegman
    - Yale University
- Eigenfaces attempt to maximise the scatter of the training images in face space.
- Fisherfaces attempt to maximise the between class scatter, while minimising the within class scatter.
- In other words, moves images of the same face closer together, while moving images of difference faces further apart.

Biometric Authentication Lecture 5 - 45

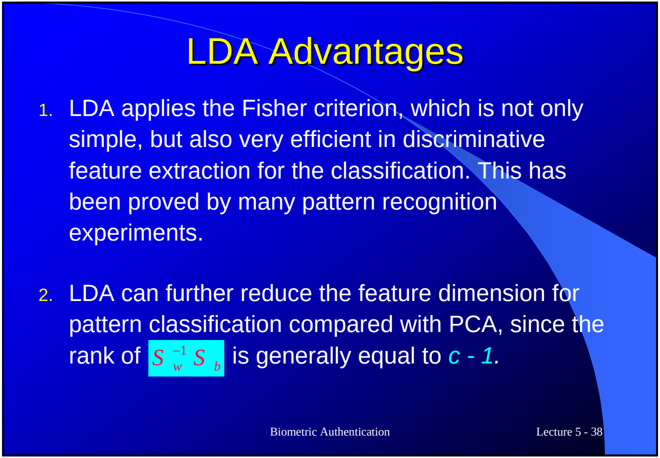
Problem 1.4 LDA

Linear discrimination analysis (LDA) is introduced in P5:37-40. Please understand the two steps in P5:37 and compare within-class scatter matrix with between-class scatter matrix.

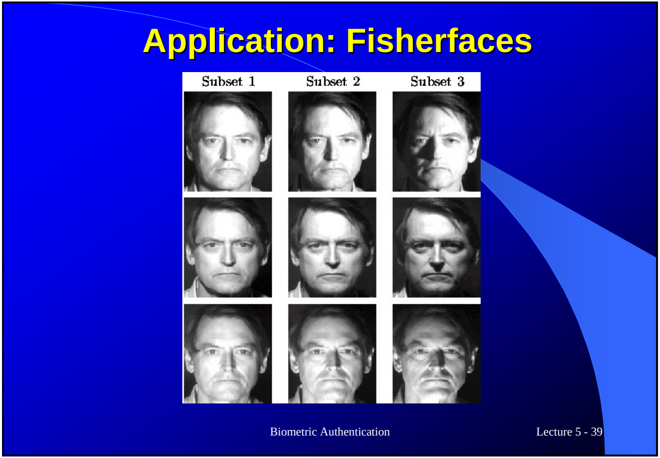
Problem 1.4 LDA



Problem 1.4 LDA



Problem 1.4 LDA



Problem 1.4 LDA

Application: Fisherfaces

Method	Subset 1	Subset 2	Subset 3
Eigenface	0.0	4.4	41.5
Eigenface w/o 1st 3	0.0	4.4	27.7
Fisherface	0.0	0.0	4.6

Biometric Authentication Lecture 5 - 40

Outline

Problem 2: Three PR Approaches

There are three PR approaches: StatPR, SyatPR and NN. What difference between them (P5:4-10)? Based on your knowledge, can you give a simple application for each approach?

Problem 2: Three PR Approaches

PR Approaches

- No single theory of PR can possibly cope with such a broad range of problems. PR applications come in many forms. A given problem may allow one or more of the different solution approaches. One of the guideline principles is just :'Use the right tool for the job!'
- Three typical PR approaches:
  1. Statistical PR (StatPR): There is an underlying and quantifiable statistical basis for the generation of patterns.
  2. Syntactic PR(SyatPR): The underlying structure of the pattern provides the information fundamental for PR.
  3. Neural Networks (NN): Neither of the above cases hold true, but we are able to develop and 'train' an architecture to correctly associate input patterns with desired responses.

Problem 2: Three PR Approaches

Statistical PR Approach (StatPR)

- StatPR - Statistical (or ‘decision-theoretic’) PR – is a very popular approach in PR. So far, most of PR systems are based on this approach.
- StatPR assumes a statistical basis for classification of algorithms. A set of characteristic measurements, denoted features, are extracted from the input data and are used to assign each feature vector to one of the *c* classes. Features are assumed generated by a state of nature, and therefore the underlying model is of a state of nature or class-conditioned set of probabilities and/or probability density functions.

Biometric Authentication　　Lecture 5 - 5

Problem 2: Three PR Approaches

Syntactic PR Approach (SyntPR)

- Many times the significant information in a pattern is not merely in the presence or absence, or the numerical values, of a set of features. Rather , the interrelationships or interconnections of features yield important structural information, which facilitates structural description or classification. This is the basis of syntactic (or structural) PR. However, in using SyntPR approaches, we must be able to quantify and extract structural information and to assess structural similarity of patterns.
- Typically, SyntPR approaches formulate hierarchical descriptions of complex patterns built up from simpler sub-patterns. At the lowest level, primitive elements or 'building blocks' are extracted from the input data.

Biometric Authentication Lecture 5 - 6

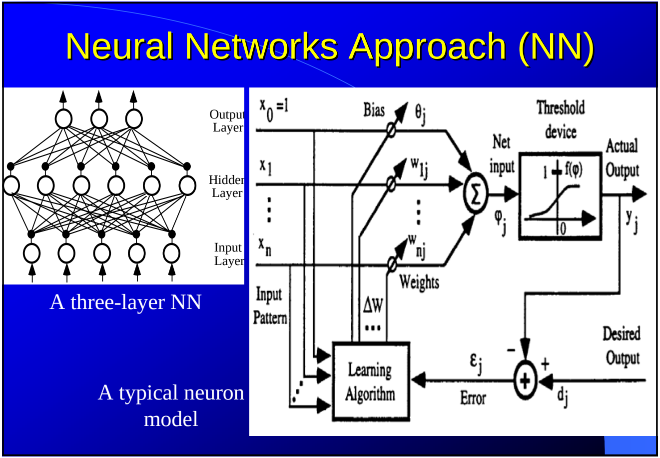
Problem 2: Three PR Approaches

Neural Networks Approach (NN)

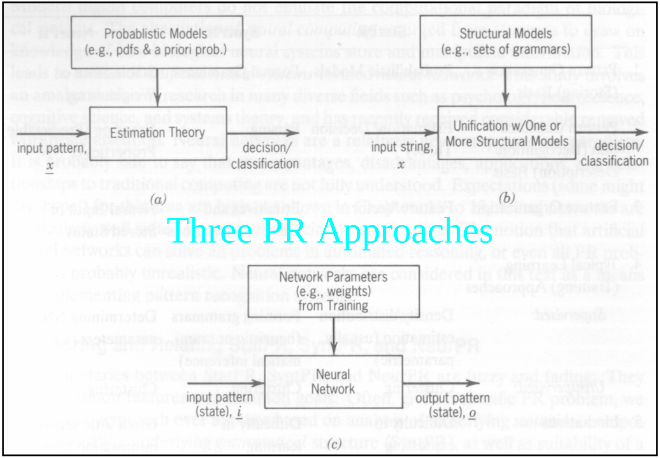
- NN is motivated by a desire to try both to understand the brain and to emulate some of its strengths. It can be characterized by
  - ∠ its pattern of connections (architectures)
  - ∠ its method of determining the weights (training or learning)
  - ∠ its activation function
- Developed as generalizations of mathematical models of human cognition or neural biology, based on the assumptions that
  - # Information processing occurs at neurons
  - # Signals are passed between neurons
  - # Each connection link has an associated weight
  - # An activation function is applied to each neurons

Biometric Authentication   Lecture 5 - 7

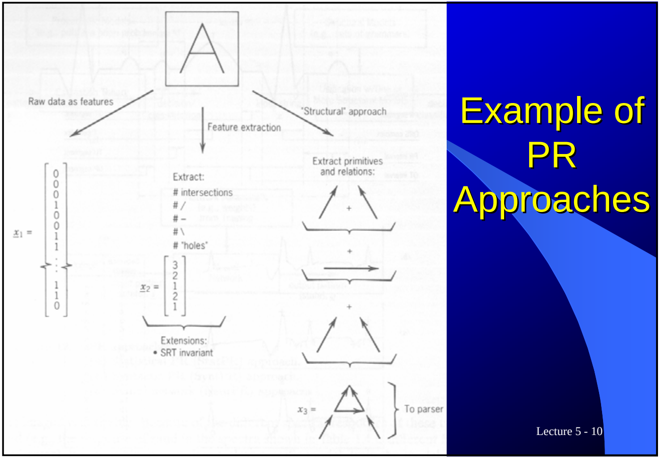
Problem 2: Three PR Approaches



Problem 2: Three PR Approaches



Problem 2: Three PR Approaches



Problem 2: Three PR Approaches

- The properties of the patterns (entities) are different:
- For statistical PR the patterns can be extracted a set of features and there is an underlying statistical model for the generation of theses patterns.
- For syntactic PR the patterns should contain structural or relational information and the structure of an entity is paramount.
- For NN, the patterns are stored in the Network, and it can't be seen.

Problem 2: Three PR Approaches

- The methods are different:
- For statistical PR the problems are solved by using statistical theory.
- For syntactic PR the classification can be accomplished by defining suitable and distinct grammars that reflect the structure of each pattern class.
- For NN, the result is learned by the neurons in the network.

Problem 2: Three PR Approaches

Examples

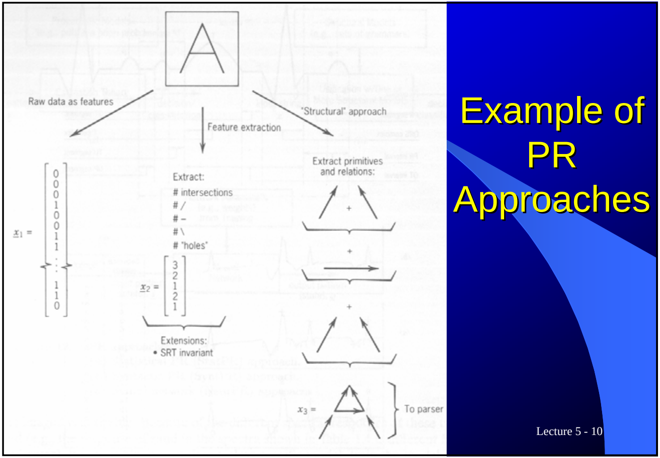
Approaches	Examples
StatPR	Face, Palmprint
SyatPR	Disease
NN	Motor Control

Outline

Problem 3: StatPR and SyatPR

Please check the examples of PR approaches in P5:10. Try to analyze the character "H" by statistical PR approach and structural PR approach.

Problem 3: StatPR and SyatPR



Problem 3: StatPR and SyatPR

StatPR approach

The feature set:(intersections, -, |, holes)

$x = [2, 1, 4, 0]$

SyntPR approach

Primitives and relations:$\uparrow + \uparrow \rightrightarrows \rightarrow \rightrightarrows \uparrow + \uparrow$

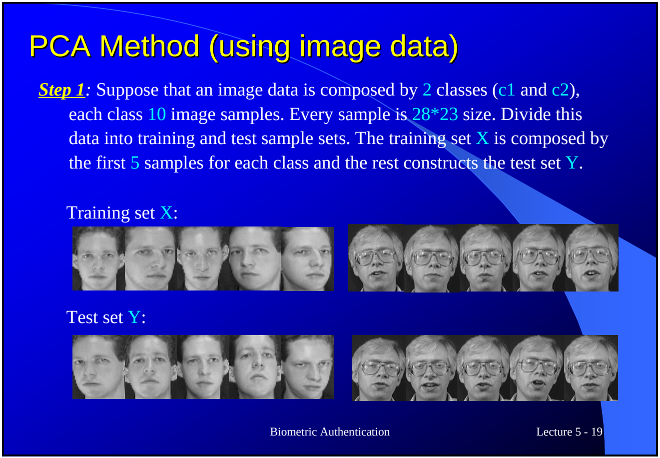
$\square$

Outline

Problem 4: PCA

From P5:19-25, PCA method given by using image data is defined, which projects an image space with 644 dimension into 6 dimension eigenvector space. Please understand each step.

Problem 4: PCA



Problem 4: PCA

PCA Method (using image data)

Step 2: Express every image sample in X, $X_j^{(i)} \; (i=1,2\,;\; j=1,\cdots,5)$,

by a 644 (=28*23) dimensional feature vector. For X, assume that its total mean value is m, the mean values of c1 and c2 are m1 and m2, *E is mathematical expectation* we have:

$$m_1 = E\left(X^{(c_1)}\right) = \sum_{j=1}^{5} p_j^{(1)} X_j^{(1)} = \sum_{j=1}^{5} \frac{1}{5} X_j^{(1)}$$ ->

$$m_2 = E\left(X^{(c_2)}\right) = \sum_{j=1}^{5} p_j^{(2)} X_j^{(2)} = \sum_{j=1}^{5} \frac{1}{5} X_j^{(2)}$$ ->

$$m = E(m_1, m_2) = \sum_{i=1}^{2} p^{(i)} m_i = \sum_{i=1}^{2} \frac{1}{2} m_i$$ ->

Problem 4: PCA

PCA Method (using image data)

Step 3: Compute the total scatter matrix S_t:

$$S_t = E\left[(X-m)(X-m)^T\right]$$

$$= \sum_{i=1}^{2} p_i \left[\sum_{j=1}^{5} p_j^{(i)} \left(X_j^{(i)}-m\right)\left(X_j^{(i)}-m\right)^T\right]$$

$$= \sum_{i=1}^{2} \frac{1}{2} \left[\sum_{j=1}^{5} \frac{1}{5} \left(X_j^{(i)}-m\right)\left(X_j^{(i)}-m\right)^T\right]$$

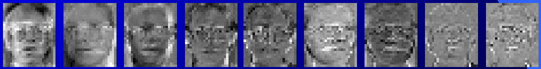
Problem 4: PCA

PCA Method (using image data)

Step 4: Compute the eigenvalue and eigenvectors of S_t We have

$$\lambda_i \varphi_i = S_t \varphi_i \quad , \quad i = 1, \cdots, 644 \quad ,$$

where λ_i and φ_i are the i^{th} eigenvalue and eigenvector. Each vector is 644 dimensional. We re-express 9 eigenvectors with nonzero eigenvalues in the form of image as following (they are also called Eigenfaces):



Select some eigenvectors that have the largest eigenvalues to represent the most significant variation within the image set

Biometric Authentication Lecture 5 - 22

Problem 4: PCA

PCA Method (using image data)

Step 5: Select the most principal components or eigenvectors. From the following figure, we find that

$$\lambda_{10} = \lambda_{11} = \cdots = \lambda_{644} = 0 \quad ,$$

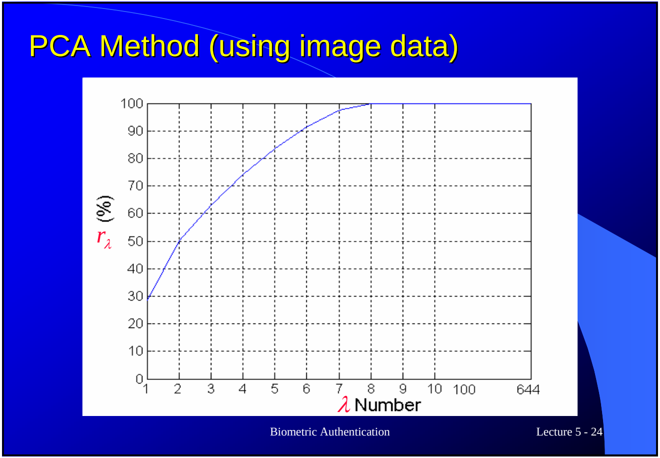
and

$$r_\lambda = \frac{\sum_{i=1}^{6} \lambda_i}{\sum_{j=1}^{644} \lambda_j} \times 100\% = 91.32\% > 90\% \quad ,$$

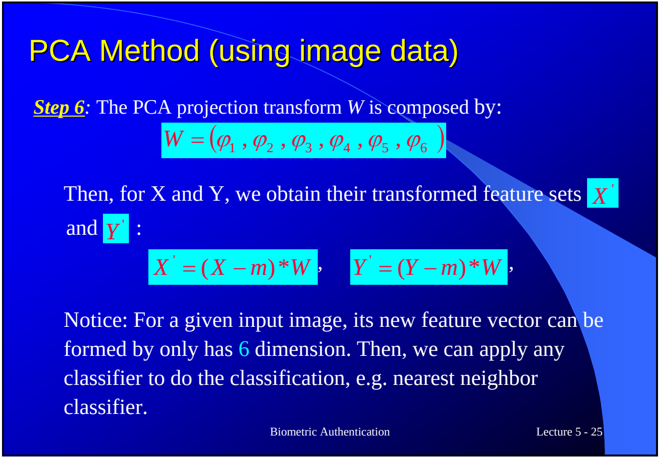
where r_λ is the ratio of the eigenvalue sum of selected components to the total sum (refer to the following figure).

Biometric Authentication Lecture 5 - 23

Problem 4: PCA



Problem 4: PCA

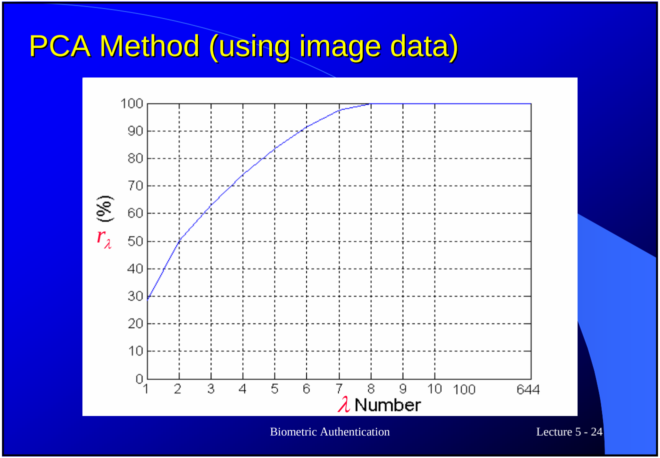


Outline

Problem 5: Three PR Approaches

According to the figure in P5:24, how to find its minimum λ *if we hope to get* $\gamma_\lambda > 60$*?*

Problem 5: Three PR Approaches



When $\lambda = 2$ the $\gamma_\lambda < 60$When $\lambda = 3$ the $\gamma_\lambda > 60$, so $\lambda = 3$ is what we want.

Problems

Any questions?