パターン認識
Pattern Recognition

佐藤真一
Shin'ichi Satoh

国立情報学研究所
National Institute of Informatics

June 13, 2023

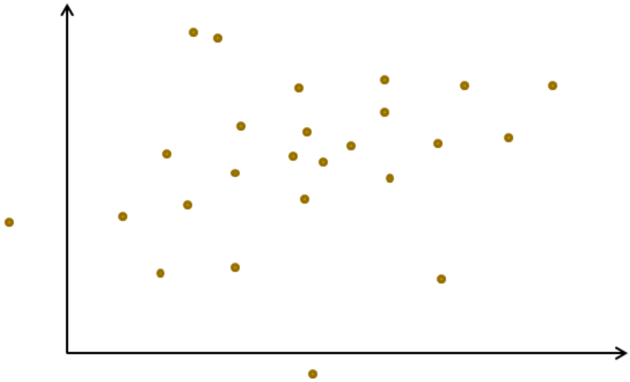
最終レポート

- いわゆるトップジャーナル・トップ国際会議の任意のパターン認識に関する論文についてまとめる
- e.g., IEEE TPAMI, IJCV, CVPR, ICCV, NeurIPS, ICML, ACMMM...
- 以下について記述せよ:
  - 論文の書誌事項
  - 論文の概要
  - どのような問題を扱っているのか、なぜその問題は重要なのか、どのように解決しているのか、どのように評価しているか
  - なぜその論文を選んだのか、何が特に興味深いのか
  - 講義への意見・コメント
- A4 2-4 ページ程度 PDF
- 締め切り: 07/31/2023
- ITC-LMS より提出

本日の予定

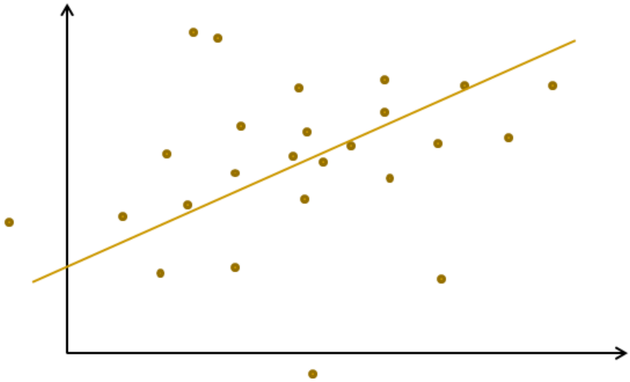
- 線形変換のパターン認識における応用例について紹介
- 顔検出と認識
  - 主成分分析 (Principal Component Analysis: PCA) と Eigenface 法
  - 線形判別分析 (Linear Discriminant Analysis: LDA) と Fisherface 法

導入



- いくつかの点がベクトル空間中に存在するとする
- 主成分分析は、これらの点群を最もよく近似する線形多様体を得る強力な手法
- いかに計算するか？どのような応用があるか？

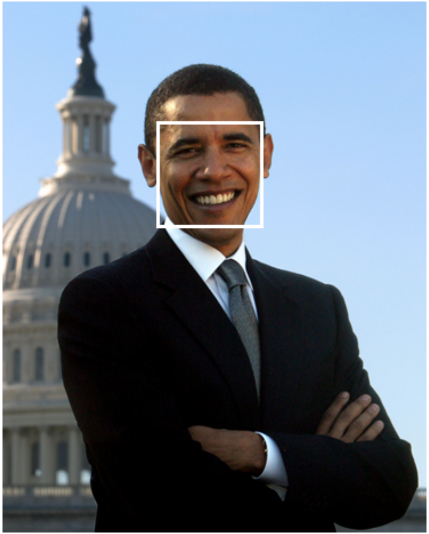
導入



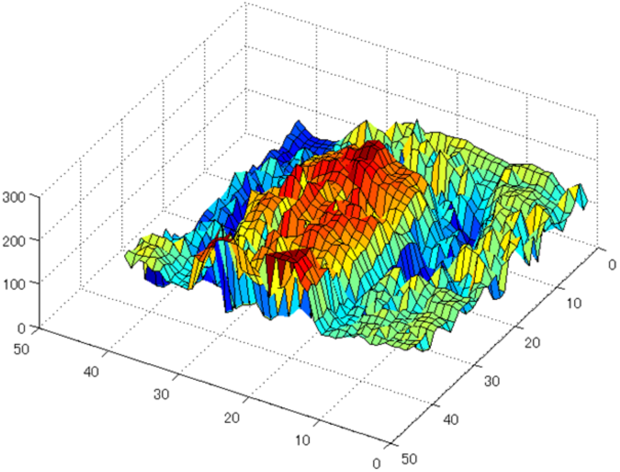
- いくつかの点がベクトル空間中に存在するとする
- 主成分分析は、これらの点群を最もよく近似する線形多様体を得る強力な手法
- いかに計算するか？どのような応用があるか？

顔検出と照合

- 一群の顔画像があるとしよう (名前との対応付き)
- 顔検出: 与えられた未知の画像が顔であるか否かを判別する
- 顔照合: 与えられた顔画像の名前を言い当てる



輝度データとしての画像



画像変換

元画像

白黒画像

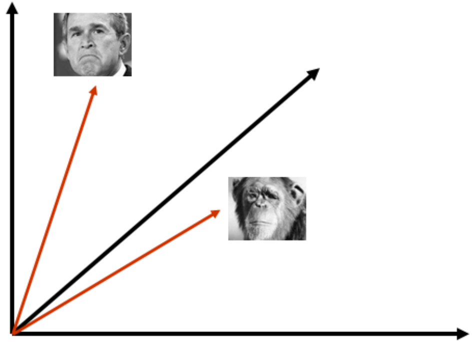
切り取り (crop)

離散化 (discretize)

量子化 (quantize)

$$\mathbf{x} = [x_1 \; x_2 \; x_3 \; \cdots \; x_n]$$

顔画像空間

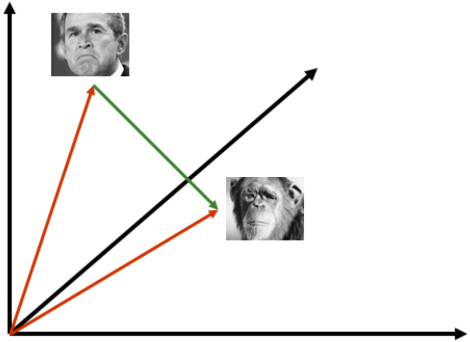


画像は高次元空間中の点として捉えられる

- $N \times M$ 画素の画像は $\mathbb{R}^{NM}$ 上の点
- 二次元空間のように、この空間上でベクトルが定義可能

[Thanks to Chuck Dyer, Steve Seitz, Nishino]

顔画像空間

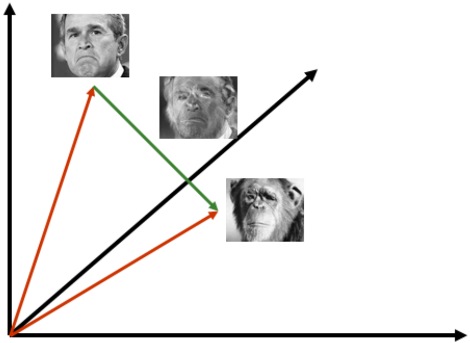


画像は高次元空間中の点として捉えられる

- $N \times M$ 画素の画像は $\mathbb{R}^{NM}$ 上の点
- 二次元空間のように、この空間上でベクトルが定義可能

[Thanks to Chuck Dyer, Steve Seitz, Nishino]

顔画像空間

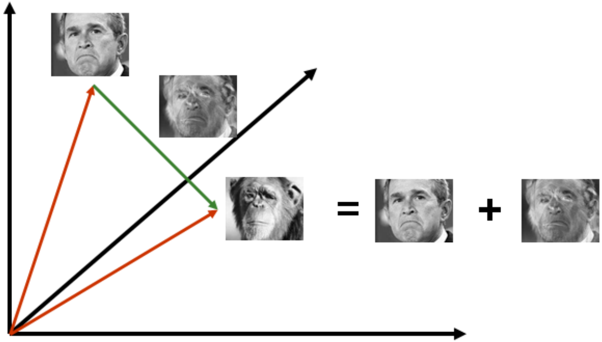


画像は高次元空間中の点として捉えられる

- $N \times M$ 画素の画像は $\mathbb{R}^{NM}$ 上の点
- 二次元空間のように、この空間上でベクトルが定義可能

[Thanks to Chuck Dyer, Steve Seitz, Nishino]

顔画像空間



画像は高次元空間中の点として捉えられる

- $N \times M$ 画素の画像は $\mathbb{R}^{NM}$ 上の点
- 二次元空間のように、この空間上でベクトルが定義可能

[Thanks to Chuck Dyer, Steve Seitz, Nishino]

正規分布

- 分布が未知の場合、通常正規分布を仮定
- $p(x|\text{non-face})$ は一様分布，$p(x|\text{face})$ は正規分布に従うと仮定しよう
- 多次元正規分布

$$N_x(M,\Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} exp(-\frac{1}{2}(X-M)^T\Sigma^{-1}(X-M))$$

正規分布

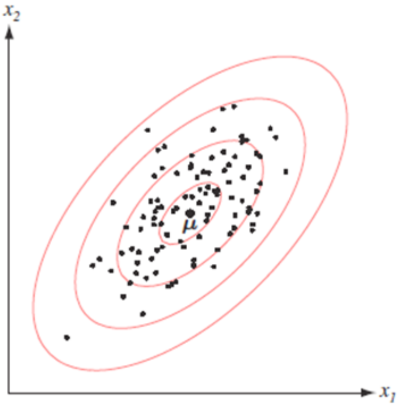


FIGURE 2.9. Samples drawn from a two-dimensional Gaussian lie in a cloud centered on the mean $\boldsymbol{\mu}$. The ellipses show lines of equal probability density of the Gaussian. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

固有顔法

- キーアイデア：顔画像の分布を画像空間中で観測
- 主成分を主成分分析 (PCA) により発見
- 未知の画像を主成分で張られる部分空間上に投影
- 顔検出：この空間に十分近ければ顔、さもなければ非顔
- 顔認識：投影した空間中で、より近い点に対応する人物と認識

固有顔法

- 顔サンプルが与えられたものとする:

$$X = [x_1\ x_2\ x_3\ \cdots\ x_n]$$

- 共分散行列を計算:

$$\mu = E[\mathbf{x}] = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\Sigma = E[(\mathbf{x}-\mu)(\mathbf{x}-\mu)^T] = \frac{1}{n}(X-\mu)(X-\mu)^T$$

固有顔法

- 共分散行列に対し固有値展開を行う

$$\Sigma\phi_i = \lambda_i\phi_i$$

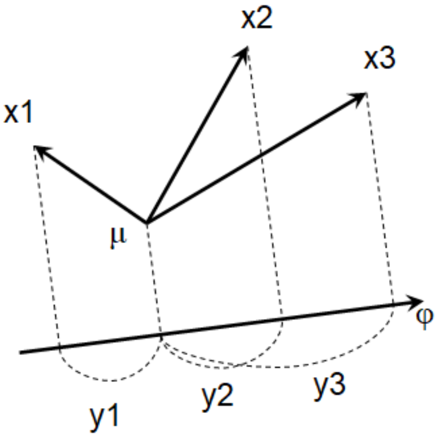
- ここで λ_i は固有値、 ϕ_i は固有ベクトル
- m 個の大きい固有値に対応する m 個の固有ベクトル $(\phi_i\ ,\ i = 1, \cdots, m)$ を求める
- これに基づき、各画像 x は m 次元ベクトルに変換する

$$\begin{aligned}\Phi &= [\phi_1\ \phi_2\ \cdots\ \phi_m] \\ x' &= \Phi^T(x-\mu)\end{aligned}$$

共分散行列の代数的・幾何的解釈

二次形式 $\phi^T \Sigma \phi$ の意味とは?

$$
\begin{aligned}
\phi^T \Sigma \phi =& \phi^T \left[\frac{1}{n} \sum_i (x_i - \mu)(x_i - \mu)^T \right] \phi \\
=& \frac{1}{n} \sum_i \left[\phi^T (x_i - \mu)(x_i - \mu)^T \phi \right] \\
=& \frac{1}{n} \sum_i \left[\phi^T (x_i - \mu) \right] \left[(x_i - \mu)^T \phi \right] \\
=& \frac{1}{n} \sum_i \left[\phi^T (x_i - \mu) \right] \left[\phi^T (x_i - \mu) \right]^T \\
=& \frac{1}{n} \sum_i \left[\phi^T (x_i - \mu) \right]^2 \\
=& \frac{1}{n} y_i^2
\end{aligned}
$$

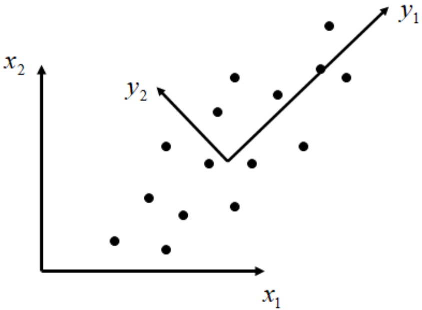


共分散行列の代数的・幾何的解釈

$\phi^T\phi = 1$ を条件として $\phi^T\Sigma\phi$ を最大化する

$$
\begin{aligned}
J =& \phi^T\Sigma\phi - \lambda(\phi^T\phi - 1) \quad \text{ラグランジュ未定乗数決定法} \\
\frac{\partial J}{\partial \phi} =& \cdots \\
=& 2\Sigma\phi - 2\lambda\phi = 0 \\
\Sigma\phi =& \lambda\phi
\end{aligned}
$$

Eigenface 法による顔検出・認識

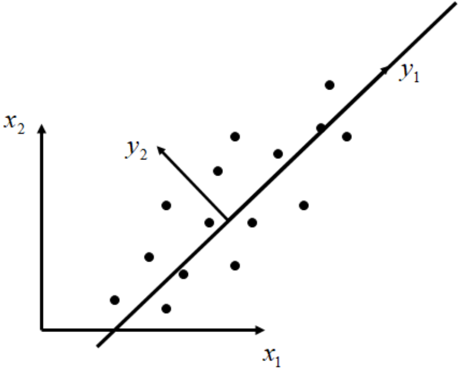


$$
\begin{aligned}
\Sigma &= E\mathbf{x}\mathbf{x}^T \\
\Sigma\phi_i &= \lambda_i\phi_i \\
& \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \\
\Phi &= [\phi_1\ \phi_2\ \cdots\ \phi_m] \quad m \ll n \\
y &= \Phi^T x \\
e &= x - \Phi y
\end{aligned}
$$

主成分分析 (PCA) とは、最小自乗誤差を実現する、確率ベクトル $\mathbf{x}$ の低次元部分空間 Y への正規直交射影

J. M. Rehg ©2002

Eigenface 法による顔検出・認識

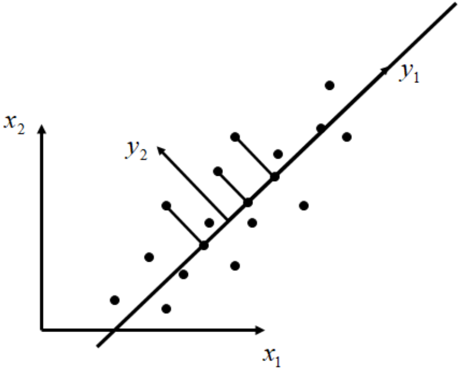


$$
\begin{aligned}
\Sigma &= E\mathbf{x}\mathbf{x}^T \\
\Sigma\phi_i &= \lambda_i\phi_i \\
& \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \\
\Phi &= [\phi_1\ \phi_2\ \cdots\ \phi_m] \quad m \ll n \\
y &= \Phi^T x \\
e &= x - \Phi y
\end{aligned}
$$

主成分分析 (PCA) とは、最小自乗誤差を実現する、確率ベクトル $\mathbf{x}$ の低次元部分空間 Y への正規直交射影

J. M. Rehg ©2002

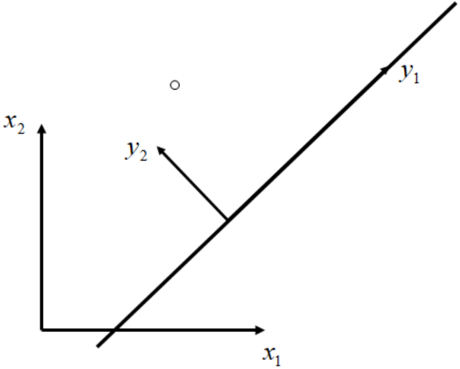
Eigenface 法による顔検出・認識



$$
\begin{aligned}
\Sigma &= E\mathbf{x}\mathbf{x}^T \\
\Sigma\phi_i &= \lambda_i\phi_i \\
& \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \\
\Phi &= [\phi_1\ \phi_2\ \cdots\ \phi_m] \quad m \ll n \\
y &= \Phi^T x \\
e &= x - \Phi y
\end{aligned}
$$

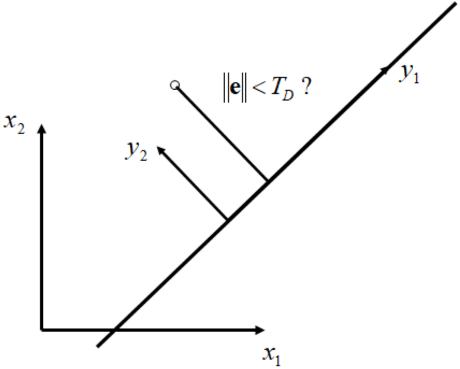
主成分分析 (PCA) とは、最小自乗誤差を実現する、確率ベクトル $\mathbf{x}$ の低次元部分空間 Y への正規直交射影

Eigenface 法による顔検出・認識



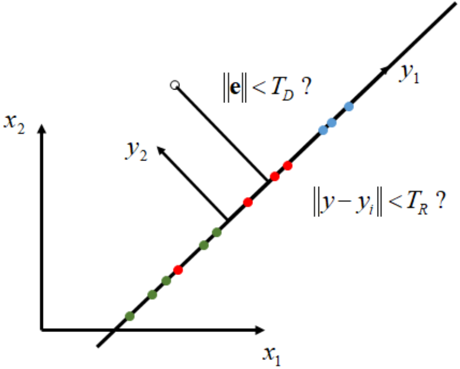
$$
\begin{aligned}
\Sigma =& E\mathbf{x}\mathbf{x}^T \\
\Sigma\phi_i =& \lambda_i\phi_i \\
& \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \\
\Phi =& [\phi_1\ \phi_2\ \cdots\ \phi_m] \quad m \ll n \\
y =& \Phi^T x \\
e =& x - \Phi y
\end{aligned}
$$

Eigenface 法による顔検出・認識



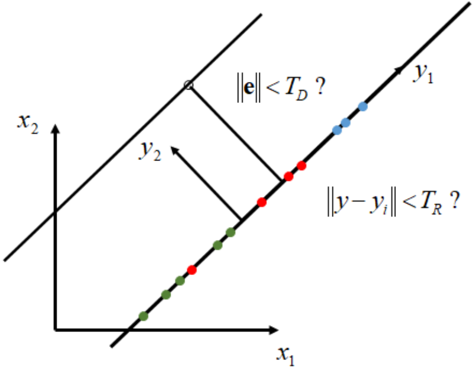
$$
\begin{aligned}
\Sigma &= E\mathbf{x}\mathbf{x}^T \\
\Sigma\phi_i &= \lambda_i\phi_i \\
&\quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \\
\Phi &= [\phi_1\ \phi_2\ \cdots\ \phi_m] \quad m \ll n \\
y &= \Phi^T x \\
e &= x - \Phi y
\end{aligned}
$$

Eigenface 法による顔検出・認識



$$
\begin{aligned}
\Sigma &= E\mathbf{x}\mathbf{x}^T \\
\Sigma\phi_i &= \lambda_i\phi_i \\
&\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \\
\Phi &= [\phi_1\ \phi_2\ \cdots\ \phi_m] \quad m \ll n \\
y &= \Phi^T x \\
e &= x - \Phi y
\end{aligned}
$$

Eigenface 法による顔検出・認識



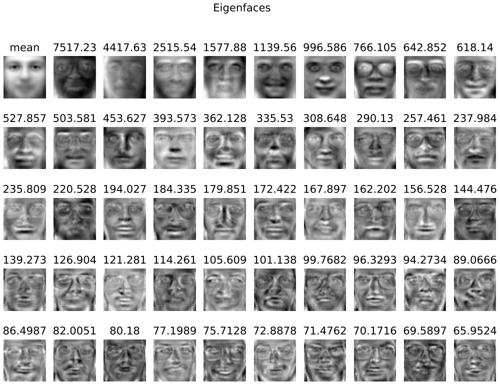
$$
\begin{aligned}
\Sigma &= E\mathbf{x}\mathbf{x}^T \\
\Sigma\phi_i &= \lambda_i\phi_i \\
& \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \\
\Phi &= [\phi_1\ \phi_2\ \cdots\ \phi_m] \quad m \ll n \\
y &= \Phi^T x \\
e &= x - \Phi y
\end{aligned}
$$

- 顔検出は顔空間からの距離で決定
- 顔認識は顔空間内での距離で決定

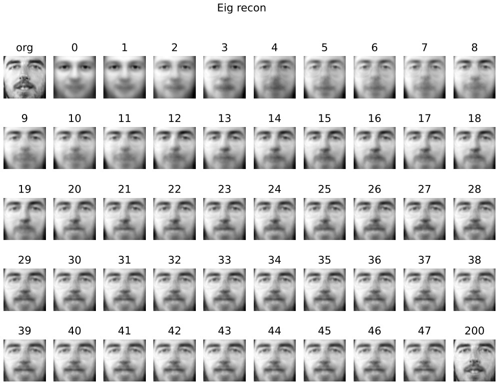
固有顔 (Eigenface) 法

PCA により固有ベクトルを求める

- $\phi_1, \phi_2, \phi_3, \cdots$ とする
- 各ベクトルは顔空間の軸の方向を規定
- 画像として可視化してみると。。。



固有顔空間への射影



固有顔 $\phi_1, \phi_2, \phi_3, \cdots, \phi_m$ により顔空間が張られる

未知の顔は以下のように固有顔軸に変換される

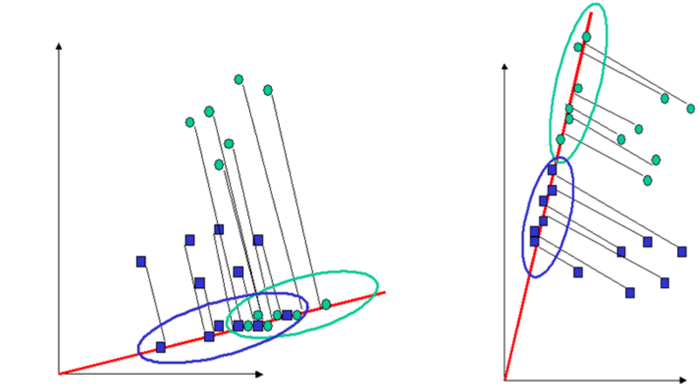
$$x \rightarrow [\underbrace{(x-\bar{x})^T\phi_1}_{y_1}\ \underbrace{(x-\bar{x})^T\phi_2}_{y_2}\ \cdots\ \underbrace{(x-\bar{x})^T\phi_m}_{y_m}]$$

$$x \approx \bar{x} + y_1\phi_1 + y_2\phi_2 + \cdots + y_m\phi_m$$

Fisherface 法

- 固有顔法は主成分分析 (PCA) により顔画像の分布を最大化する軸 (主成分) を見つけていた
- もし顔の人物名情報も使ったらどうか？
- 異なる人物の顔画像間の分散を最大化し、同じ人物の顔画像間の分散を最小化するのは理にかなった考え方のように思われる
- 線形判別分析 (Linear Discriminant Analysis: LDA) によりそうした軸を見つけることができる
- こうして発見された軸で張られた空間に画像を射影する
- 顔認識 (のみ): 射影された空間上でもっとも近い顔画像の人物に識別する

Fisherface 法



Poor Projection　　　　Good Projection

Fisherface 法

- N 個の顔画像

$$\{x_1, x_2, \cdots, x_n\}$$

- C クラスに分類されている

$$\{\chi_1, \chi_2, \cdots, \chi_C\}$$

- 各クラスの平均

$$\mu_i = \frac{1}{|\chi_i|} \sum_{x \in \chi_i} x$$

- 全体の平均

$$\mu = \frac{1}{N} \sum_{k=1}^{N} x_k$$

Fisherface 法

- 各クラス i の分散:

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

- クラス内分散:

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

- クラス間分散:

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

- 全分散

$$S_T = S_W + S_B$$

Fisherface 法

- 入力 x を低次元ベクトル y に変換する線形射影 W を求めたい:

$$y = W^T x$$

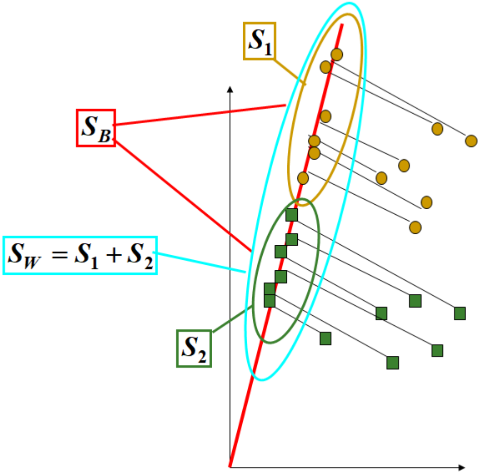
- 空間 y におけるクラス間分散:

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

- 空間 y におけるクラス内分散:

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

Fisherface 法



Good separation

Fisherface 法

- 求めたい射影は:

$$W_{opt} = \arg\max_{W} \frac{|\tilde{S_B}|}{|\tilde{S_W}|} = \arg\max_{W} \frac{|W^T S_B W|}{|W^T S_W W|}$$

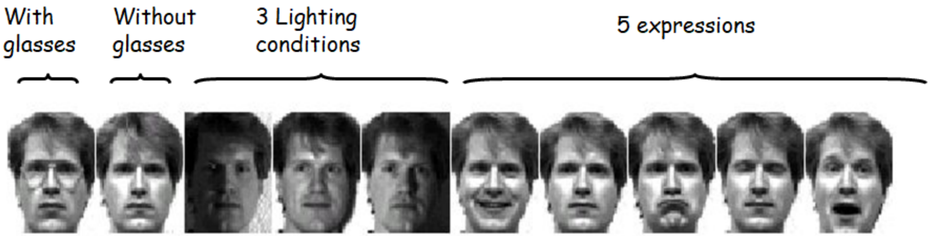
- これは一般固有値展開により求めることができる

$$S_B w_i = \lambda_i S_W w_i \quad i = 1, \cdots, m$$

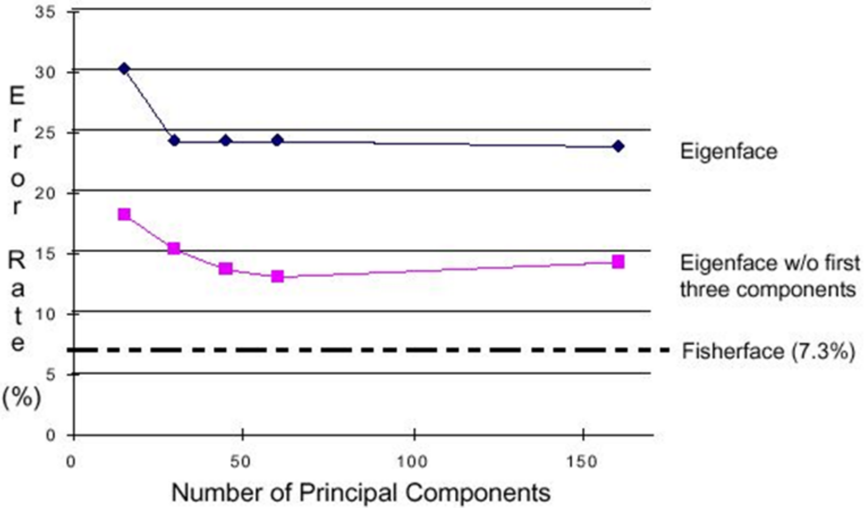
Fisherface 法: 実験

表情、眼鏡有無、照明条件等の変動のある顔画像データセット Yale を用いて実験

- 16 人分 160 画像
- 学習: 159 画像、評価: 1 画像 (Leave-one-out 法)



Fisherface 法: 実験



デモ

- 顔データを用いて Eigenfaces と Fisherfaces の実験
- Eigenfaces 可視化
- Eigenfaces による顔画像の再合成
- Eigenfaces と Fisherfaces の認識率比較
- Olivetti faces dataset を利用
- 40 人、各 10 画像 (照明条件、表情など変化)