

基于 PCA 与最大后验概率分类的人脸识别方法

袁少锋 王士同

(江南大学数字媒体学院 无锡 214122)

摘 要 在运用主成分分析进行人脸识别的过程中,由于实际图像可能符合某种概率密度分布,并且实际用到的图像可能受到不同程度的噪声污染,简单的距离分类已不再适用。基于核函数的最大后验概率分类是将概率密度函数估计中的参数估计、核函数以及贝叶斯理论结合起来,能很好地考虑到概率分布情况,用多元高斯分布下的基于核函数的最大后验概率分类取代距离分类,对于含有不同参数值的高斯噪声图像有较好的识别率。用 ORL 标准人脸库进行验证,实验结果表明了可行性。

关键词 主成分分析,多元高斯分布,参数估计,核函数,贝叶斯理论

中图分类号 TP391.41 文献标识码 A

Method of Face Recognition Based on Principal Component Analysis and Maximum a Posteriori Probability Classification

YUAN Shao-feng WANG Shi-tong

(School of Digital Media, Jiangnan University, Wuxi 214122, China)

Abstract In the processing of face recognition with PCA algorithm, the image may be eligible for some kind of probability density distribution and different levels of noise pollution, so the simple distance classification is no longer effective. Maximum posteriori classification combines the parameter estimation and kernel function and Bayes theory, can take into account the probability distribution well. Under the multivariate Gaussian distribution, using it to replace the distance classification can have the better recognition rate for the images containing the different parameter values of the Gaussian noise. The standard ORL face library was used to verify this theory, and the result shows its feasibility.

Keywords Principal component analysis, Multivariate gaussian distribution, Parameter estimation, Kernel function, Bayesian theory

1 引言

近年来,人脸识别已成为人工智能、模式识别等学科的研究热点,是既具有理论价值又有应用价值的重要的研究课题。在此期间,Turk 和 Pentland^[1]提出特征脸(Eigenface)方法,并将主成分分析^[2](Principal Component Analysis, PCA)算法应用到人脸识别领域中,经过不断发展与完善,PCA 方法已成为人脸识别的基准算法,目前被广泛应用。然而 PCA 在进行分类识别中,采用的是距离分类,常用的距离分类有基于欧氏距离的最近邻分类、基于哈曼顿距离的三阶近邻分类^[3]等。距离分类虽然是简单而有效的分类方法,但是并没有考虑图像的实际概率密度分布。并且,现实所用到的图像大都受到不同程度的污染,而这种污染也可以近似估计为符合某种概率密度分布^[4](比如:高斯分布)。简单的距离分类不能很好地考虑这些因素,经过查找,目前也没有文献明确给出合理的分类来解决这类问题,本文将基于核函数的最大后验概率分类^[5]取代距离分类,对此类问题有很好的改善。从理论上分析,将基于核函数的最大后验概率分类引入 PCA(记为 N_PCA)进行人脸识别,与传统的基于距离分类的 PCA(记为 O

_PCA)进行比较,N_PCA 方法会对含有噪声(本文为高斯噪声)的图像有比较好的识别率。本文用 ORL 标准人脸库进行验证,实验结果表明了 N_PCA 的可行性。

针对上面表述,以下章节将会对距离分类和基于核函数的最大后验概率分类进行简要阐述。

2 距离分类

用 $L(x, y)$ 表示样本 x 与 y 之间的距离,样本的维数为 k ,则有以下常见的距离函数。

(1) Minkowsky 距离:距离公式可表示为:

$$L(x, y) = \left[\sum_{i=1}^k |x_i - y_i|^\lambda \right]^{\frac{1}{\lambda}} \quad (1)$$

(2) 最近邻法:最近邻法的距离为欧几里得距离或欧几里得距离的开平方根,它是 Minkowsky 距离当 $\lambda=2$ 时的特例。距离公式可表示为:

$$L(x, y) = \left[\sum_{i=1}^k |x_i - y_i|^2 \right]^{\frac{1}{2}} \quad (2)$$

计算出最小距离的两幅图像属于同一类别。

(3) 三阶近邻法:三阶近邻法的距离为曼哈顿距离,它是 Minkowsky 距离当 $\lambda=1$ 时的特例。距离公式可表示为:

到稿日期:2013-05-20 返修日期:2013-07-16

袁少锋(1988-),男,硕士生,主要研究方向为人工智能与模式识别;王士同(1964-),男,教授,主要研究方向为模式识别、人工智能、生物信息。

$$L(x, y) = \sum_{i=1}^k |x_i - y_i| \quad (3)$$

三阶近邻需要计算出与测试图像距离最小的 3 幅图像^[6], 分别将这 3 幅图像所属的类记为 class1、class2、class3, 若 class1 和 class2 且 class2 和 class3 不属于同一类别, 则测试图像属于 class1; 若 class1 和 class2 属于同一类别, 则测试图像属于 class1; 若 class2 和 class3 属于同一类别, 则测试图像属于 class2。

3 基于核函数的最大后验概率分类

3.1 核函数

目前对于核函数的构造有多种方法, 一个函数能够成为核函数, 必须满足 Mercer 条件。对于给定的任何样本, 其核函数必然是存在的, 并且选择合适的核函数是影响分类的关键因素。核函数发展至今, 在实际应用中被广泛应用的核函数 $k(x_i, x_j)$ 主要是以下几种:

(1) 高斯径向核函数 (RBF 核函数):

$$k(x_i, x_j) = \exp(-\|x_i - x_j\|_2^2 / \sigma^2) \quad (4)$$

式中, σ 为尺度参数, 在很大程度上影响着 RBF 核函数的性能。

(2) 多项式核函数:

$$k(x_i, x_j) = (\langle x_i - x_j \rangle + c)^d \quad (5)$$

式中, c 为常数, d 为多项式阶数。当 $c=0, d=1$ 时, 多项式核函数变为线性核函数。

(3) Sigmoid 核函数:

$$k(x_i, x_j) = \tanh(\text{scale} \times \langle x_i - x_j \rangle - \text{offset}) \quad (6)$$

式中, scale 和 offset 分别是尺度和衰减参数。

人脸图像受到高斯噪声污染后, 其特征脸的特征分布与原始相比, 发生了较大变化。由于 RBF 核函数作为一种局部性核函数, 能够根据局部特征比较好地进行平滑运算, 分类性能好, 故本文采用 RBF 核函数。

3.2 类条件概率密度函数

假设有 N 个独立同分布的 d 维数据 $\{x_1, x_2, \dots, x_N\}$, 分为 $\{C_1, C_2, \dots, C_m\}$ 共 m 类数据, 多元高斯分布为 $N(u, \Sigma)$, 即:

$$p(x) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \times \exp[-\frac{1}{2}(x-u)^T \Sigma^{-1}(x-u)] \quad (7)$$

式中, u 是中心, Σ 是对称、正定矩阵。

将利用 PCA 所得到的特征人脸通过非线性映射函数 $\Phi(\cdot)$ 映射到新的特征空间, 在新的特征空间中, 多元高斯分布的类条件概率密度函数为:

$$p(\Phi(x) | C_i) = \frac{1}{(2\pi)^{N/2} |\Sigma_i|^{1/2}} \times \exp[-\frac{1}{2}(\Phi(x) - u_i)^T \Sigma_i^{-1}(\Phi(x) - u_i)] \quad (8)$$

式中, N 为空间维数。

将式(8)两边取对数并去掉常数项, 得到

$$g_i(\Phi(x)) = (\Phi(x) - u_i)^T \Sigma_i^{-1}(\Phi(x) - u_i) + \ln |\Sigma_i| \quad (9)$$

可以将式(8)变形为:

$$p(\Phi(x) | C_i) = \frac{1}{(2\pi)^{N/2}} \exp[-\frac{1}{2}g_i(\Phi(x))] \quad (10)$$

确立类的条件概率密度函数 $p(\Phi(x) | C_i)$ 后, 根据贝叶斯理论^[9]有:

$$p(C_i | \Phi(x)) = \frac{p(\Phi(x) | C_i) p(C_i)}{\sum_{i=1}^m p(\Phi(x) | C_i) p(C_i)} \quad (11)$$

如果有

$$p(C_\omega | \Phi(x)) = \max_{1 \leq j \leq m} [p(C_j | \Phi(x))] \quad (12)$$

则 $x \in C_\omega$ 。

3.3 参数估计

采用最大似然估计法^[8]可得到均值 u_i 和协方差矩阵 Σ_i 的表达式如下:

$$u_i = \frac{1}{N_i} \sum_{j=1}^{N_i} [\Phi(x_j)] \quad (13)$$

$$\Sigma_i = S_i \quad (14)$$

$$S_i = \frac{1}{N_i} \sum_{j=1}^{N_i} [\Phi(x_j) - u_i][\Phi(x_j) - u_i]^T \quad (15)$$

在 $\Phi(\cdot)$ 映射的空间中, 如果直接用协方差矩阵 Σ_i , 将出现二次项的判别函数, 会增加问题的复杂性。所以, 利用式(16)对协方差矩阵 Σ_i 进行规整:

$$\Sigma_i(\theta, \eta) = (1 - \eta) \Sigma_i(\theta) + \eta \frac{\text{trace}(\Sigma_i(\theta))}{p} I \quad (16)$$

式中, $\theta(0 \leq \theta \leq 1)$, $\eta(0 \leq \eta \leq 1)$, I 为单位矩阵, p 为训练样本数, $S = \sum_{i=1}^m S_i$, $\Sigma_i(\theta) = (1 - \theta) S_i + \theta S$ 。

3.4 公式的核化

根据线性代数知识可知, Σ_i 是正定、对称矩阵, 可以用以下公式进行对角化:

$$\Sigma_i = \sum_{j=1}^N \Lambda_{ij} v_{ij} v_{ij}^T \quad (17)$$

并且有: $|\Sigma_i| = \prod_{j=1}^N \Lambda_{ij}$ 。

其中, Λ_{ij} 和 v_{ij} 分别为协方差矩阵 Σ_i 的特征值与特征向量。

由于比较小的特征值对结果影响比较小, 因此本文选用前 k 个比较大的特征值, 并且用第 $k+1$ 个特征值 h_i 代表第 $k+1$ 以后的所有特征值。

此时式(9)可表示为:

$$\begin{aligned} g_i(\Phi(x)) &= \sum_{j=1}^k \frac{1}{\Lambda_{ij}} [v_{ij}^T(\Phi(x) - \mu_i)]^2 + \sum_{j=k+1}^N \frac{1}{h_i} [v_{ij}^T(\Phi(x) - \mu_i)]^2 + \ln(h_i^{N-k} \prod_{j=1}^k \Lambda_{ij}) \\ &= \frac{1}{h_i} (\sum_{j=1}^k [v_{ij}^T(\Phi(x) - \mu_i)]^2 - \sum_{j=1}^k (1 - \frac{h_i}{\Lambda_{ij}}) [v_{ij}^T(\Phi(x) - \mu_i)]^2) + \ln(h_i^{N-k} \prod_{j=1}^k \Lambda_{ij}) \end{aligned} \quad (18)$$

令:

$$g_{1i}(\Phi(x)) = \sum_{j=1}^k [v_{ij}^T(\Phi(x) - \mu_i)]^2 \quad (19)$$

$$g_{2i}(\Phi(x)) = \sum_{j=1}^k (1 - \frac{h_i}{\Lambda_{ij}}) [v_{ij}^T(\Phi(x) - \mu_i)]^2 \quad (20)$$

则:

$$g_i(\Phi(x)) = \frac{1}{h_i} \{g_{1i}(\Phi(x)) - g_{2i}(\Phi(x))\} + \ln(h_i^{N-k} \prod_{j=1}^k \Lambda_{ij}) \quad (21)$$

由线性代数知识, v_{ij} 可表示为:

$$v_{ij} = \sum_{l=1}^N \gamma_{ij}^{(l)} \Phi(x_l) = U \gamma_{ij} \quad (22)$$

式中, $\gamma_{ij} = (\gamma_{ij}^{(1)}, \dots, \gamma_{ij}^{(N)})^T$, $U = (\Phi(x_1), \dots, \Phi(x_N))$ 为 $\Phi(\cdot)$ 映射空间的一组标准基。

将上式代入得:

$$g_{ii}(\Phi(x)) = \sum_{j=1}^N [\gamma_{ij}^T U^T (\Phi(x) - \mu_i)^T (\Phi(x) - \mu_i) U \gamma_{ij}]$$

$$= \sum_{j=1}^N [\gamma_{ij}^T (k_x - \frac{1}{N_i} \sum_{l=1}^{N_i} k_{xl}) (k_x - \frac{1}{N_i} \sum_{l=1}^{N_i} k_{xl})^T \gamma_{ij}] \quad (23)$$

$$g_{zi}(\Phi(x)) = \sum_{j=1}^k \{ (1 - \frac{h_i}{\lambda_{ij}}) \times [\gamma_{ij}^T (k_x - \frac{1}{N_i} \sum_{l=1}^{N_i} k_{xl}) (k_x - \frac{1}{N_i} \sum_{l=1}^{N_i} k_{xl})^T \gamma_{ij}] \} \quad (24)$$

式中,

$$k_x = (k(x_1, x), k(x_2, x), \dots, k(x_N, x))^T$$

$$k_{xl} = (k(x_1, x_l), k(x_2, x_l), \dots, k(x_N, x_l))^T$$

由于 Λ_{ij} 与 v_{ij} 分别是 Σ_i 的特征值与特征向量, 因此:

$$v_{ij}^T \Sigma_i v_{ij} = \Lambda_{ij} \quad (25)$$

即:

$$\gamma_{ij}^T U^T \{ \frac{1}{N_i} \sum_{j=1}^{N_i} [\Phi(x_j) - u_i] [\Phi(x_j) - u_i]^T \} U \gamma_{ij}$$

$$= \gamma_{ij}^T \frac{1}{N_i} \sum_{j=1}^{N_i} (k_{xj} - \frac{1}{N_i} \sum_{l=1}^{N_i} k_{xl}) (k_{xj} - \frac{1}{N_i} \sum_{l=1}^{N_i} k_{xl})^T \gamma_{ij}$$

$$= \Lambda_{ij}$$

上式可知, γ_{ij} 是 $\gamma_{ij}^T \frac{1}{N_i} \sum_{j=1}^{N_i} (k_{xj} - \frac{1}{N_i} \sum_{l=1}^{N_i} k_{xl}) (k_{xj} - \frac{1}{N_i} \sum_{l=1}^{N_i} k_{xl})^T \gamma_{ij}$ 的特征向量。

4 实验研究与分析

为了说明用 PCA 进行人脸识别, 基于核函数的最大后验概率分类好于简单的距离分类, 本节选取标准的 ORL 人脸库图像, 并对这些图像添加不同程度的高斯噪声, 然后分别采用 O_PCA 与 N_PCA 方法进行实验, 在同等条件下, 将两组实验结果进行比较。

4.1 人脸数据集

ORL 人脸数据库包含 40 个人, 每人有 10 幅图像。图像为单一深色背景的正面图像, 包含了一定的光照变化、表情变化(睁眼与闭眼、笑与不笑)、面部细节变化(戴眼镜与不戴眼镜)以及一定范围内的深度旋转。这些图像的大小均为 112×92 像素。我们选取其中的 20 人, 每人 10 幅, 共 200 张图像作为实验样本。图 1 显示了部分样本人脸图像。



图 1 ORL 部分样本人脸图像

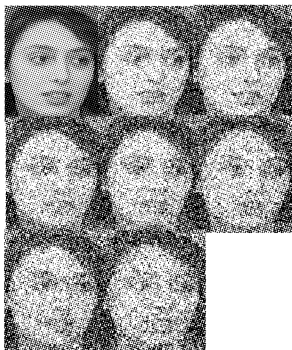


图 2 ORL 受污染程度不同的某人脸图像

由于在实际应用中含有不同类型的噪声, 本文将噪声模型估计为高斯噪声, 向样本集中添加高斯噪声, 设置不同的高斯噪声参数值, 从而得到同种噪声模型下受污染程度不同的人脸图像, 添加噪声后的某人脸图像如图 2 所示。图 2 中, 从左到右分别为未受污染的某人脸图像, 参数值均值 $\mu=0.1$ 、方差分别为 $0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1$ 的受污染图像。

4.2 实验结果及分析

分别运用 O_PCA 和 N_PCA 进行人脸识别, 文献[6]表明, 三阶近邻法比其他距离分类算法有更高的识别率。则 O_PCA 中的距离分类选用基于哈曼顿距离的三阶近邻分类, 实验中需要用到的参数如表 1 所列。

表 1 实验参数

名称	参数	数值
RBF 核函数	σ	$\sigma \in (1, 100)$
$\Sigma_i(\theta, \eta)$	θ, η	$\theta \in (0, 1), \eta \in (0, 1)$

为了讨论参数对实验结果的影响, 以每人的随机 2 幅图像为训练样本, 剩余的 8 幅图像为测试样本, 只改变其中一个参数值, 其他参数固定。实验结果为 10 次实验的平均值, 如图 3、图 4 所示。

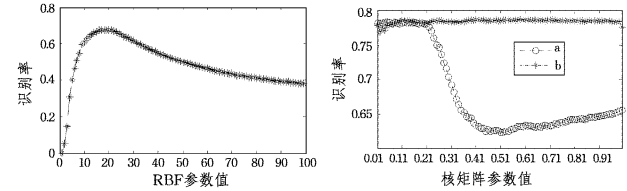


图 3 RBF 参数对结果的影响 图 4 参数 $\theta=b, \eta=a$ 对结果的影响

从图 3 可以看出, RBF 参数 $\sigma=15$ 附近有较好的实验值; 从图 4 可以看出, $\Sigma_i(\theta, \eta)$ 的参数在 $\theta \in [0.1, 0.99], \eta \in [0.01, 0.2]$ 中取值会有较好效果。由于文章篇幅的限制, 只针对表 2 所列的参数取值进行实验。

表 2 参数及参数值

名称	参数	数值
RBF 核函数	σ	13, 15, 17
$\Sigma_i(\theta, \eta)$	θ, η	$\theta=0.1, \eta=0.03$
高斯噪声	μ, σ^2	$\mu=0.1, \sigma^2 \in [0.04, 0.1]$

在实验中, 随机选取每人的 2、3、4、5、6 幅图像作为训练样本, 对应剩下的 8、7、6、5、4 作为测试样本。实验结果如表 3 所列, 实验结果为 10 次实验的平均值, \circ 里的值为 10 次结果的方差。

从表 3 可以看到, 随着训练样本数增加, 两种算法的识别率均有提高。在样本数相同的前提下, N_PCA 算法的识别效果要好于 O_PCA; RBF 核函数参数的取值对 N_PCA 的结果有一定的影响, 合理科学地确定 RBF 核函数参数值会有更好的实验结果。

选取表 3 中噪声参数为 $\sigma^2=0.08, \sigma^2=0.1$, RBF 核函数参数值为 15、每类随机样本数为 6 时的 10 次运行结果, 如图 5、图 6 所示。

从图 5、图 6 可以看出, 在参数值相同的情况下, 几乎每次运行结果 N_PCA 都要好于 O_PCA。这是因为传统的距离分类没有考虑到图像的实际概率密度分布问题, 而基于核函数的最大后验概率分类结合了概率统计分布知识, 能够比较合理地估计图像的实际概率密度分布。

表 3 含不同程度的高斯噪声的人脸识别率(%)

噪声参数 σ^2	算法	人脸识别率(%)					
		训练样本数					
		2	3	4	5	6	
0.04	O_PCA	76.81(± 0.12)	82.43(± 0.19)	88.50(± 0.13)	91.00(± 0.15)	92.13(± 0.08)	
	N_PCA	RBF 参数 13	78.94(± 0.20)	86.14(± 0.26)	90.75(± 0.06)	92.30(± 0.05)	94.50(± 0.10)
		15	79.69(± 0.14)	85.57(± 0.20)	90.67(± 0.12)	92.70(± 0.16)	95.87(± 0.06)
		17	79.50(± 0.24)	86.57(± 0.12)	90.58(± 0.10)	91.90(± 0.11)	94.75(± 0.13)
0.05	O_PCA	72.56(± 0.31)	80.86(± 0.14)	85.92(± 0.11)	90.20(± 0.10)	91.38(± 0.06)	
	N_PCA	RBF 参数 13	79.13(± 0.34)	86.00(± 0.05)	90.75(± 0.20)	92.70(± 0.06)	93.87(± 0.07)
		15	79.31(± 0.16)	85.36(± 0.15)	89.67(± 0.11)	92.60(± 0.07)	93.88(± 0.06)
		17	80.00(± 0.16)	85.64(± 0.36)	89.25(± 0.13)	93.20(± 0.10)	93.75(± 0.11)
0.06	O_PCA	64.69(± 0.11)	77.07(± 0.39)	83.83(± 0.20)	90.60(± 0.15)	91.62(± 0.06)	
	N_PCA	RBF 参数 13	75.44(± 0.48)	83.57(± 0.14)	87.33(± 0.32)	91.80(± 0.08)	93.38(± 0.05)
		15	75.50(± 0.12)	82.13(± 0.13)	86.33(± 0.16)	91.20(± 0.08)	93.50(± 0.06)
		17	75.69(± 0.42)	82.00(± 0.08)	87.67(± 0.24)	92.40(± 0.12)	93.63(± 0.04)
0.07	O_PCA	61.75(± 0.64)	73.79(± 0.26)	84.58(± 0.26)	89.40(± 0.04)	90.75(± 0.05)	
	N_PCA	RBF 参数 13	65.63(± 0.27)	82.93(± 0.24)	90.33(± 0.06)	93.30(± 0.06)	93.87(± 0.03)
		15	72.81(± 0.26)	83.26(± 0.27)	87.67(± 0.24)	92.40(± 0.12)	93.37(± 0.10)
		17	70.25(± 0.64)	82.21(± 0.23)	87.92(± 0.29)	90.70(± 0.10)	92.87(± 0.12)
0.08	O_PCA	57.87(± 0.38)	72.36(± 0.08)	79.33(± 0.25)	87.50(± 0.14)	90.12(± 0.06)	
	N_PCA	RBF 参数 13	69.00(± 0.70)	83.79(± 0.16)	84.25(± 0.16)	90.90(± 0.12)	93.25(± 0.11)
		15	69.44(± 0.73)	83.36(± 0.12)	84.50(± 0.15)	90.90(± 0.09)	93.25(± 0.12)
		17	72.69(± 0.26)	83.71(± 0.21)	85.00(± 0.12)	90.20(± 0.09)	92.83(± 0.07)
0.09	O_PCA	54.81(± 0.11)	71.79(± 0.31)	80.13(± 0.51)	83.10(± 0.17)	88.13(± 0.12)	
	N_PCA	RBF 参数 13	68.37(± 0.64)	80.64(± 0.21)	87.25(± 0.28)	88.90(± 0.13)	91.87(± 0.11)
		15	68.75(± 0.85)	80.71(± 0.22)	84.67(± 0.23)	88.70(± 0.22)	91.63(± 0.16)
		17	68.75(± 0.28)	79.79(± 0.33)	82.33(± 0.29)	88.10(± 0.16)	91.25(± 0.12)
0.1	O_PCA	49.50(± 0.42)	67.21(± 0.91)	77.67(± 0.37)	81.00(± 0.51)	87.50(± 0.17)	
	N_PCA	RBF 参数 13	66.19(± 0.49)	82.57(± 0.30)	84.58(± 0.34)	87.20(± 0.22)	90.88(± 0.19)
		15	66.44(± 0.41)	82.79(± 0.49)	84.33(± 0.12)	86.40(± 0.16)	90.65(± 0.11)
		17	66.81(± 0.34)	83.17(± 0.27)	83.67(± 0.12)	87.70(± 0.14)	90.83(± 0.13)

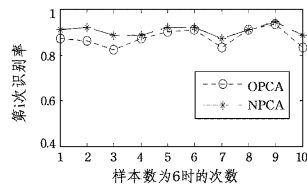
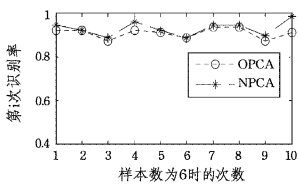


图 5 $\sigma^2=0.08$ 时的 10 次运行结果 图 6 $\sigma^2=0.1$ 时的 10 次运行结果

算法 O_PCA 与 N_PCA 进行比较:

(1)识别精度:当人脸图像含有较多高斯噪声点时,N_PCA 比 O_PCA 的识别率高很多,因此,N_PCA 比 O_PCA 识别效果好。

(2)抗噪性:高斯噪声方差从 0.04~0.1,N_PCA 的识别效果都要好于 O_PCA,说明 N_PCA 有较强的抗噪性。

(3)复杂性:由于 N_PCA 在识别过程中有许多参数需要调试,比 O_PCA 耗时,但有比较好的实验结果。

结束语 运用 PCA 进行人脸识别,用多元高斯分布下基于核函数的最大后验概率分类取代距离分类,N_PCA 与 O_PCA 相比,N_PCA 对含有高斯噪声的图像有很好的识别率,但是,N_PCA 有参数需要调整,调参数比较耗时。

参 考 文 献

[1] Turk M, Pentland A. Eigenfaces for recognition[J]. Journal of

Cognitive Neuroscience, 1993, 3(1): 71-86
 [2] Tan Ke-ren, Chen Song-can. Adaptively weighted sub-pattern PCA for face recognition[J]. Neurocomputing, 2005, 64(3): 505-511
 [3] Jain A K, Robert P W, et al. Statistical Pattern Recognition: A Review [J]. IEEE Transactions on pattern analysis and machine intelligence, 2000, 22(1): 4-37
 [4] He Kun, Luan Xin-cheng, et al. Gaussian Noise Removal of Image on the Local Feature[J]. Intelligent Information Technology Application, 2008, 3(1): 867-871
 [5] Xu Zeng-lin, Huang Kai-zhu, Zhu Jian-ke, et al. A novel kernel-based maximum a posteriori classification method [J]. Neural Networks, 2009, 22(7): 977-987
 [6] 张俭鸽, 刘洪波. WTPCA 和三阶近邻的人脸识别算法仿真[J]. 计算机工程与应用, 2009, 45(11): 175-177
 [7] Motai Y, Yoshida h. Principal Composite Kernel Feature Analysis: Data-Dependent Kernel Approach[J]. IEEE Transactions on Knowledge and Date Engineering, 2012(99): 1-13
 [8] 茆诗松, 程依明, 濮晓龙. 概率论与数理统计教程[M]. 北京: 高等教育出版社, 2004: 287-290
 [9] Lukic A S, Wernick M N. Bayesian Kernel Methods for Analysis of Functional Neuroimages[J]. IEEE Transactions on Medical Imaging, 2007, 26(12): 1613-1624