

PCA 与 2DPCA 的关系

闫荣华^{1,2} 彭进业^{1,3} 汶德胜²

(西北工业大学电子信息学院 西安 710072)¹ (中国科学院西安光学精密机械研究所 西安 710119)²
(西北大学信息科学与技术学院 西安 710127)³

摘要 主成分分析和二维主成分分析是两种经典的数据变换方法。尽管许多学者对 PCA 和 2DPCA 进行了大量的研究和实验,但并未给出 PCA 与 2DPCA 之间的联系。文中给出二者之间的联系,即 PCA 与 2DPCA 在优化时具有相同的最优目标值,同时通过理论推导和在 CMU-PIE 与 CK+库上的实验证明了这一观点。

关键词 人脸识别,主成分分析,二维主成分分析

中图分类号 TP751 **文献标识码** A

Relationship between PCA and 2DPCA

YAN Rong-hua^{1,2} PENG Jin-ye^{1,3} WEN De-sheng²

(School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China)¹

(Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China)²

(School of Information and Technology, Northwest University, Xi'an 710127, China)³

Abstract Principal component analysis and two-dimensional principal component analysis are two classical data transformation techniques. Although many scholars worked on them, there's no relationship of two methods to be given. In this paper, we gave their relationship, which is that they have the same optimal objective value in optimization. The opinion was demonstrated through theoretical derivation and extensive experiments which is tested on the CMU-PIE and CK+ face database.

Keywords Face recognition, PCA, 2DPCA

1 引言

近 30 年来,人脸识别技术得到了迅速发展,涌现出了许多新的方法。其中主成分分析法(PCA)无疑是人脸识别中里程碑式的数据变换方法。

PCA 方法^[1]于 1901 年由 Karl Pearson 首次提出。Sirovich 和 Kirby^[2]将 PCA 应用于识别,之后 Matthew Turk 和 Alex Pentland^[3-4]将其应用于人脸识别领域。这种简单有效的方法是基于 Karhunen-Loeve(KL)^[5]变换的。其基本思想是:首先将人脸图像转换成高维向量,然后对高维向量进行 KL 变换,从而得到特征脸,特征脸可以作为特征子空间的一个基,最后将人脸图像投影到这个基上,并得到相应的低维特征向量。但是,由于 PCA 方法需要将人脸图像转换成一维向量,导致原始人脸图像的结构信息被破坏,会引起维数灾难。

针对以上问题,许多学者提出了改进方法,其中以 2004 年杨健等人^[6-7]提出的二维主成份分析(2DPCA)最为突出,与 PCA 方法基于向量不同,2DPCA 方法基于 2D 图像矩阵。因此,在 2DPCA 中构建协方差矩阵时,不需要事先将图像矩阵变换成向量。图像协方差矩阵直接由原始样本图像矩阵构建,进而得到用于特征提取的特征向量。与 PCA 方法构建的

协方差阵的规模相比,2DPCA 方法的图像协方差阵要小得多,因此,杨健等人认为与 PCA 方法相比,2DPCA 方法具有 3 个优点:1)图像协方差阵估计准确;2)图像协方差阵的特征分解速度快;3)在所有实验中 2DPCA 的识别错误率低于 PCA。

王立伟认为 2DPCA 与分块 PCA 等价^[8],2DPCA 是分块 PCA 的一个特例。但是,GAO 指出王立伟的等价性证明不够严密^[9],只有当输入图像被变换成零均值时才成立,而对于一般情形,等价性是不成立的。但是其并未给出 PCA 与 2DPCA 的联系。

本文将给出 PCA 与 2DPCA 的相互联系,并通过理论证明和大量的实验证明它们在优化时具有相同的最优目标值。

本文第 2 节分析 PCA 方法和 2DPCA 方法;第 3 节给出理论证明;第 4 节给出实验结果;最后总结全文。

2 PCA 和 2DPCA

本节将分别介绍 PCA 与 2DPCA 方法的共同点和各自的特点。

2.1 PCA

PCA 方法通过 KL 变换消除了原有向量的各分量之间的相关性,从而去掉了带有较少信息的坐标轴,以达到降低特

本文受基于图像的植物识别和检索研究(61272285),国家 863 课题,媒体大数据的结构化描述方法研究(2014AA015201)资助。

闫荣华(1976—),男,博士生,主要研究方向为机器学习、深度学习,E-mail: yanronghua@opt. ac. cn;彭进业(1964—),男,博士,教授,主要研究方向为图像检索、机器学习;汶德胜(1964—),男,博士,教授,主要研究方向为高光谱图像处理。

征空间维数的目的。

设有 N 幅样本图像集合 $\{A_k\}$, 其中 $k=1, 2, \dots, N$, A_k 为 $m \times n$ 的样本图像矩阵。将 A_k 列向量化得到 x_k , x_k 是 mn 维的向量。对人脸做线性变换, 将 mn 维的图象空间映射到 q 维的特征空间 ($q \leq mn$), 则特征向量 $y_k \in R^q$ 由式(11)给出。

$$y_k = W^T x_k, k=1, 2, \dots, N \quad (1)$$

其中, $W \in R^{mn \times q}$ 为一个各列相互单位正交的矩阵。

为了得到最优的投影矩阵 W , 采用样本的总分散度测量投影矩阵 W 的好坏。样本的总分散度可由特征空间的协方差阵的迹定义, 即:

$$J(W) = \text{tr}(S_y) \quad (2)$$

其中, S_y 为特征空间的协方差阵, $\text{tr}(S_y)$ 为 S_y 的迹。通过使式(2)中的准则最大化可以得到最优投影矩阵 W , 当所有样本投影到最优投影矩阵 W 上时, 特征空间中映射后样本的总分散度达到最大。

特征空间的协方差阵 S_y 可以定义为:

$$\begin{aligned} S_y &= E(Y - EY)(Y - EY)^T \\ &= E[(W^T X - E(W^T X))(W^T X - E(W^T X))^T] \\ &= E[W^T(X - EX)][(X - EX)^T W] \end{aligned}$$

因此,

$$\text{tr}(S_y) = \text{tr}[W^T E(X - EX)(X - EX)^T W] \quad (3)$$

如果样本总离差阵 S_T 定义为:

$$\begin{aligned} S_T &= E(X - EX)(X - EX)^T \\ &= \frac{1}{N} \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T \end{aligned} \quad (4)$$

其中, $\mu \in R^{mn}$ 为所有样本的均值图像, 则在式(2)中的准则可以表示为:

$$J(W) = \text{tr}(W^T S_T W) \quad (5)$$

在 PCA 中, 对 S_T 进行特征分解, 选择与最大的 q 个特征值相对应的特征向量作为投影方向 W_{opt} 。它们满足单位正交约束和使 $J(W)$ 达到最大的条件。

$$\{w_1, \dots, w_q\} = \arg \max J(W)$$

$$w_i^T w_i = 1, i=1, \dots, q \quad (6)$$

$$w_i^T w_j = 0, ij, i, j=1, \dots, q$$

其中, $\{w_i | i=1, 2, \dots, q\}$ 为 S_T 的前 q 个最大特征值对应的特征向量, 每一幅人脸图像都可以投影到由 w_1, w_2, \dots, w_q 张成的子空间中。由于 w_1, w_2, \dots, w_q 这些图像很像人脸, 该方法被称为特征脸。

2.2 2DPCA

PCA 方法需要对人脸样本矩阵进行向量化, 而 2DPCA 直接对人脸矩阵进行变换, 避免了维数灾难。

与 PCA 相同, 设有 N 幅样本图像集合 $\{A_k\}$, $k=1, 2, \dots, N$, A_k 为 $m \times n$ 的样本图像矩阵。令 W 为 $n \times d$ 维各列相互单位正交的矩阵, 将 A_k 投影在 W 方向上, 使得:

$$Y_k = A_k W \quad (7)$$

其中, Y_k 为 $m \times d$ 维矩阵。

与 PCA 类似, 准则可以采用式(8):

$$J(W) = \text{tr}(S_Y) \quad (8)$$

其中, S_Y 为特征空间的协方差阵, $\text{tr}(S_Y)$ 为 S_Y 的迹。通过使式(8)中的准则最大化可以得到最优投影矩阵 W , 当所有样

本投影到最优投影矩阵 W 上时, 特征空间中映射后样本的总分散度达到最大。

将特征空间的协方差阵 S_Y 定义为:

$$\begin{aligned} S_Y &= E(Y - EY)(Y - EY)^T \\ &= E[(AW - E(AW))(AW - E(AW))^T] \\ &= E[(A - EA)W][(A - EA)W]^T \\ &= E[(A - EA)W][W^T(A - EA)^T] \end{aligned}$$

因此,

$$\text{tr}(S_Y) = \text{tr}[W^T E(A - EA)^T (A - EA)W] \quad (9)$$

如果样本总离差阵 G_T 定义为:

$$\begin{aligned} G_T &= E(A - EA)^T (A - EA) \\ &= \frac{1}{N} \sum_{k=1}^N (A_k - \bar{A})^T (A_k - \bar{A}) \end{aligned} \quad (10)$$

其中, $\bar{A} \in R^{m \times n}$ 为所有人脸样本的均值, 则在式(8)中的准则可以表示为:

$$J(W) = \text{tr}(W^T G_T W) \quad (11)$$

在 2DPCA 中对 G_T 进行特征分解, 选择与最大的 d 个特征值相对应的特征向量作为投影方向 W_{opt} 。它们满足单位正交约束和使 $J(W)$ 达到最大的条件。

$$\{w_1, \dots, w_d\} = \arg \max J(W)$$

$$w_i^T w_i = 1, i=1, \dots, d \quad (12)$$

$$w_i^T w_j = 0, ij, i, j=1, \dots, d$$

其中, $\{w_i | i=1, 2, \dots, d\}$ 为 G_T 的前 d 个最大特征值对应的特征向量, 每一幅样本图像都可以投影到由 w_1, w_2, \dots, w_d 张成的子空间中。

PCA 的主元是一个标量, 而 2DPCA 的主元是一个向量。为了公平比较两种方法的性能, 本文将主元定义为一个标量, 即 2DPCA 的一个主元向量代表的是该向量长度个主元。

2.3 两种方法的联系

PCA 的样本总离差阵反映了图像灰度值之间的相关性, 主对角线元素是每一灰度的方差, 其余元素是各灰度间的协方差。

2DPCA 的样本总离差阵反映了图像空间结构本身的相关性, 主对角线元素是每一行元素的方差, 其余元素是各行元素的协方差。

虽然 PCA 样本总类间离差阵为 $(mn)^2 \times (mn)^2$ 维, 2DPCA 样本总离差阵为 $n \times n$ 维, 但是两种方法的协方差阵的迹是相等的, 即它们在优化时具有相同的最优目标值。

3 PCA 与 2DPCA 联系的理论推导

引理 1 PCA 与 2DPCA 的协方差阵的迹相等。

证明: 设有 N 幅样本图像集合 $\{A_k\}$, $k=1, 2, \dots, N$, A_k 为 $m \times n$ 的样本图像矩阵。

$$A_k = \begin{bmatrix} a_{11k} & \dots & a_{1nk} \\ \vdots & \ddots & \vdots \\ a_{m1k} & \dots & a_{mnk} \end{bmatrix}, k=1, 2, \dots, N$$

由均值的定义得:

$$\bar{A} = \frac{1}{N} \sum_{k=1}^N A_k = \begin{bmatrix} \frac{1}{N} \sum_{k=1}^N a_{11k} & \dots & \frac{1}{N} \sum_{k=1}^N a_{1nk} \\ \vdots & \ddots & \vdots \\ \frac{1}{N} \sum_{k=1}^N a_{m1k} & \dots & \frac{1}{N} \sum_{k=1}^N a_{mnk} \end{bmatrix}$$

则

$$A_k - \bar{A} = \begin{pmatrix} a_{11k} - \frac{1}{N} \sum_{k=1}^n a_{11k} & \cdots & a_{1nk} - \frac{1}{N} \sum_{k=1}^n a_{1nk} \\ \vdots & \ddots & \vdots \\ a_{m1k} - \frac{1}{N} \sum_{k=1}^n a_{m1k} & \cdots & a_{mnk} - \frac{1}{N} \sum_{k=1}^n a_{mnk} \end{pmatrix} = \frac{1}{N} \sum_{k=1}^n \begin{pmatrix} b_{11k}^2 & \cdots & b_{11k} b_{mnk} \\ \vdots & \ddots & \vdots \\ b_{mnk} b_{11k} & \cdots & b_{mnk}^2 \end{pmatrix}$$

令 $b_{ijk} = a_{ijk} - \frac{1}{N} \sum_{k=1}^n a_{ijk}$, 其中 $i=1, 2, \dots, m; j=1, 2, \dots, n; k=1, 2, \dots, N$. 那么

$$A_k - \bar{A} = \begin{pmatrix} b_{11k} & \cdots & b_{1nk} \\ \vdots & \ddots & \vdots \\ b_{m1k} & \cdots & b_{mnk} \end{pmatrix}$$

$$(A_k - \bar{A})^T (A_k - \bar{A}) = \begin{pmatrix} b_{11k} & \cdots & b_{m1k} \\ \vdots & \ddots & \vdots \\ b_{1nk} & \cdots & b_{mnk} \end{pmatrix} \begin{pmatrix} b_{11k} & \cdots & b_{1nk} \\ \vdots & \ddots & \vdots \\ b_{m1k} & \cdots & b_{mnk} \end{pmatrix}$$

$$= \begin{pmatrix} b_{11k}^2 + \cdots + b_{m1k}^2 & \cdots & b_{11k} b_{1nk} + \cdots + b_{m1k} b_{mnk} \\ \vdots & \ddots & \vdots \\ b_{1nk} b_{11k} + \cdots + b_{mnk} b_{m1k} & \cdots & b_{1nk}^2 + \cdots + b_{mnk}^2 \end{pmatrix}$$

因此

$$G_T = \frac{1}{N} \sum_{k=1}^n (A_k - \bar{A})^T (A_k - \bar{A}) = \frac{1}{N} \sum_{k=1}^n \begin{pmatrix} b_{11k}^2 + \cdots + b_{m1k}^2 & \cdots & b_{11k} b_{1nk} + \cdots + b_{m1k} b_{mnk} \\ \vdots & \ddots & \vdots \\ b_{1nk} b_{11k} + \cdots + b_{mnk} b_{m1k} & \cdots & b_{1nk}^2 + \cdots + b_{mnk}^2 \end{pmatrix}$$

由迹的定义可以得到:

$$\text{tr}(G_T) = \frac{1}{N} \sum_{k=1}^n (b_{11k}^2 + \cdots + b_{m1k}^2 + \cdots + b_{1nk}^2 + \cdots + b_{mnk}^2) = \frac{1}{N} \sum_{k=1}^n (b_{11k}^2 + \cdots + b_{mnk}^2)$$

将 A_k 列向量化得到:

$$x_k = [a_{11k}, a_{21k}, \dots, a_{m1k}, a_{12k}, a_{22k}, \dots, a_{m2k}, a_{1nk}, a_{2nk}, \dots, a_{mnk}]^T$$

由均值的定义可得:

$$\mu = \frac{1}{N} \sum_{k=1}^n x_k = [\frac{1}{N} \sum_{k=1}^n a_{11k}, \dots, \frac{1}{N} \sum_{k=1}^n a_{m1k}, \dots, \frac{1}{N} \sum_{k=1}^n a_{mnk}]^T$$

则

$$x_k - \mu = [a_{11k} - \frac{1}{N} \sum_{k=1}^n a_{11k}, \dots, a_{m1k} - \frac{1}{N} \sum_{k=1}^n a_{m1k}]^T$$

$$\text{令 } b_{ijk} = a_{ijk} - \frac{1}{N} \sum_{k=1}^n a_{ijk}, \text{ 其中 } i=1, 2, \dots, m; j=1, 2, \dots,$$

$n; k=1, 2, \dots, N$. 那么

$$x_k - \mu = [b_{11k}, \dots, b_{m1k}, \dots, b_{mnk}]^T$$

$$(x_k - \mu)(x_k - \mu)^T = [b_{11k}, \dots, b_{m1k}, \dots, b_{mnk}]^T [b_{11k}, \dots, b_{m1k}, \dots, b_{mnk}]$$

$$= \begin{pmatrix} b_{11k}^2 & \cdots & b_{11k} b_{mnk} \\ \vdots & \ddots & \vdots \\ b_{mnk} b_{11k} & \cdots & b_{mnk}^2 \end{pmatrix}$$

因此

$$S_T = \frac{1}{N} \sum_{k=1}^n (x_k - \mu)(x_k - \mu)^T$$

由迹的定义可以得到:

$$\text{tr}(S_T) = \frac{1}{N} \sum_{k=1}^n (b_{11k}^2 + \cdots + b_{mnk}^2)$$

由于 $\text{tr}(G_T) = \text{tr}(S_T)$, 故引理得证。

引理 2 若矩阵 A 和矩阵 B 均为 $m \times m$ 矩阵, 且 B 非奇异, 则 $\text{tr}(B^{-1}AB) = \text{tr}(BAB^{-1}) = \text{tr}(A)$.

证明过程参见文献[10].

定理 PCA 与 2DPCA 在准则 $J(W) = \text{tr}(W^T S W)$ 下具有相同的最优目标值, 在 PCA 中, S 是指 S_T , W 是指 W_{PCA} ; 在 2DPCA 中, S 是指 G_T , W 是指 W_{2DPCA} .

证明: 由于 W_{PCA} 和 W_{2DPCA} 是正交阵, 因此 W_{PCA} 与 W_{2DPCA} 可逆。

由引理 2 得:

$$\text{tr}(W_{PCA}^{-1} S_T W_{PCA}) = \text{tr}(S_T)$$

$$\text{tr}(W_{2DPCA}^{-1} G_T W_{2DPCA}) = \text{tr}(G_T)$$

$$\text{又 } W_{PCA}^T = W_{PCA}^{-1}, W_{2DPCA}^T = W_{2DPCA}^{-1} :$$

$$\text{tr}(W_{PCA}^T S_T W_{PCA}) = \text{tr}(S_T)$$

$$\text{tr}(W_{2DPCA}^T G_T W_{2DPCA}) = \text{tr}(G_T)$$

又由引理 1, 得到 $\text{tr}(G_T) = \text{tr}(S_T)$, 因此 $\text{tr}(W_{PCA}^T S_T W_{PCA}) = \text{tr}(W_{2DPCA}^T G_T W_{2DPCA})$. 即 $J(W_{PCA}) = J(W_{2DPCA})$.

故 PCA 与 2DPCA 在准则 $J(W) = \text{tr}(W^T S W)$ 下具有相同的最优目标值。

4 实验结果

本节将选择两个不同的人脸数据库对 PCA 与 2DPCA 进行分析讨论, 分别从光照变化、姿态变化和表情变化对 PCA 与 2DPCA 进行全面测试。因此, 实验采用具有不同光照、姿态变化的 CMU-PIE 库和表情变化的 CK+ 库。

4.1 光照变化实验

第一个实验测试 PCA 与 2DPCA 在不同光照条件下的人脸分类识别错误率, 我们选用由卡内基梅隆大学构建的 CMU-PIE^[11] 人脸数据库。

CMU-PIE 人脸库由 68 个人的 41368 张人脸图像组成, 其中包括每个人的 13 种姿态条件、43 种光照条件和 4 种表情。图 1 为 CMU-PIE 人脸库中一个人 13 种不同姿态的图像。

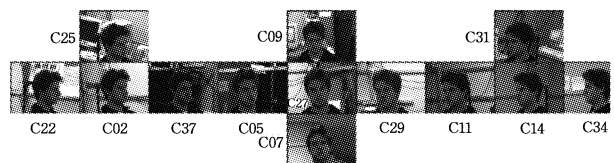


图 1 CMU-PIE 人脸库中一个人不同姿态的图像

本次实验选择了正面姿态(C27)中关闭室内灯和开启室内灯条件下的人脸集合, 共 2856 张人脸图像(每人 42 张, 一共 68 人)。图 2 为选取的关闭室内灯和开启室内灯条件下的人脸图像示例, 其中第一行为关闭室内灯条件, 第二行为开启室内灯条件。



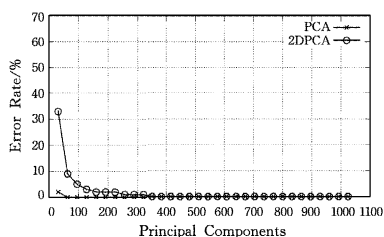
图 2 姿态 C27 下关闭室内灯和打开室内灯的光照图像对比

在选取的数据集中,每个人保持头部静止,分别用不同方向的光源照射。这里用球面角表征光源方向,球面角以大约 22.5° 的增幅增加。我们根据球面角将数据集分成 5 个子集(见表 1)。

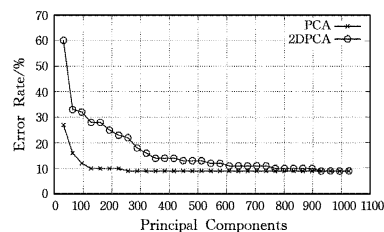
表 1 光照变化实验中按球面角分成的 5 个子集

子集	每个子集的图像数量	球面角
1	408	约 0°
2	816	约 22.5°
3	816	约 45°
4	272	约 67.5°
5	544	约 90°

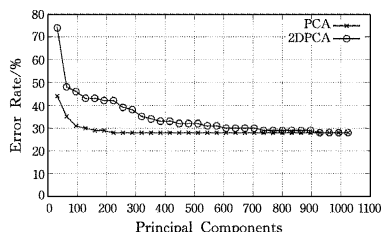
为了消除头的轮廓对识别的影响,所有图像首先被裁剪成 32×32 大小,然后采用直方图修正技术使图像具有统一的均值和方差,以提高两种方法的识别性能。实验中分类方法采用最近邻决策规则。实验采用子集 1、子集 2 作为训练集,子集 3—子集 5 作为测试集。两种方法的实验结果如图 3 所示。



(a)子集 3



(b)子集 4



(c)子集 5

图 3 两种方法在选取不同主元数时的分类错误率

以上的实验揭示了一些有趣的问题:

两种方法随着光照方向逐渐偏离正面,其错误率上升很快,这说明 PCA 与 2DPCA 对光照角度变化的鲁棒性不强。这是因为 PCA 与 2DPCA 在使样本总离差达到最大时,不仅使得对分类有用的类间离差增大,而且也使得对分类无用的类内离差同时增大。而类间和类内离差阵的迹反映了类别的可分离性,即类间离差阵的迹越大,表明类别可分离性越好;类内离差阵的迹越大,表明类别可分离性越差。从表 2 可以

看出,子集 4 的类间离差阵的迹大于子集 5 的类间离差阵的迹,说明随着光照角度逐渐偏离正面,对分类有用的类间离差减少了;而子集 4 的类内离差阵的迹小于子集 5 的类内离差阵的迹,说明对分类无用的类内离差在增加。因此随着光照方向逐渐偏离正面,两种方法错误率上升很快。

表 2 子集 4 和子集 5 的样本类内与类间离差阵的迹

方法	tr(S _b)		tr(S _w)	
	子集 4	子集 5	子集 4	子集 5
PCA	798	662	12530	27207
2DPCA	798	662	12530	27207

注:tr(S_b)为 S_b 的迹,tr(S_w)为 S_w 的迹

由图 3 可以发现,随着主元数的增加,两种方法的错误率逐渐趋于一致。尤其是当两种方法同时选取其最大主元数时,在不同的测试集上表现出相同的错误率,这前面分析的最好证明。这与主元包含的信息息息相关,主元包含的信息可以由主元的累计贡献率来表征。前 m 个主元的累计贡献率定义为 $\sum_{k=1}^m \lambda_k / \sum_{i=1}^p \lambda_i$ (m ≤ p),在 PCA 中 p=q,λ_k,λ_i 为 S_T 的特征值;在 2DPCA 中 p=d,λ_k,λ_i 为 G_T 的特征值。图 4 给出了两种方法的主元的累计贡献率趋势图。从图 4 可以看到,随着主元数的增加,两种方法的主元累计贡献率逐渐接近,它们的识别错误率同时也逐渐接近。当两种方法同时选取其最大主元数时,它们的主元累计贡献率相等,识别错误率也相等。

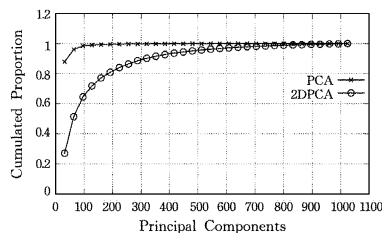


图 4 主元的累计贡献率趋势图

4.2 姿态变化实验

第二个实验仍然采用 CMU-PIE 人脸库,用于测试不同姿态对识别性能的影响。在本次实验中,我们选择关闭室内灯条件下的人脸集合,每个人有 13 种姿态,每种姿态有 3 张照片,因此每个人有 39 张人脸图像,68 人共有 2652 张人脸图像,图 5 给出了关闭室内灯条件下的 13 种姿态。根据头的位置与相机光轴的夹角将其分成 5 个子集(见表 3)。



图 5 关闭室内灯条件下的 13 种姿态

表 3 姿态变化实验的 5 个子集

子集	每个子集的图像数量	球面角
1	612	约 0°
2	408	约 22.5°
3	408	约 45°
4	816	约 67.5°
5	408	约 90°

所有人脸样本被裁剪和标准化成 32×32 大小,实验中分类方法采用最近邻决策规则。实验采用子集 1、子集 5 作为训练集,子集 2—4 作为测试集。图 6 给出了两种方法在选取不同主元数时的分类识别错误率。

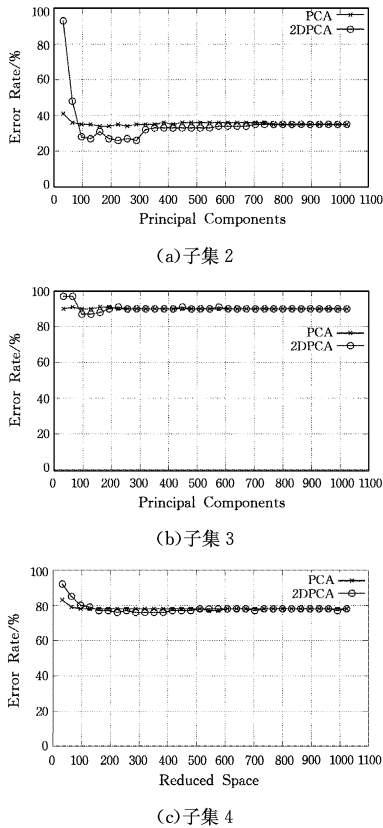


图6 两种方法在选取不同主元数时的分类识别错误率

在姿态变化实验中,可以得到以下结论:

两种方法随着人脸逐渐转向正面时,错误率上升很快,这说明PCA与2DPCA对姿态变化的鲁棒性不强。这是因为PCA与2DPCA是线性降维方法,而人脸姿态变化是非线性变化的。两种方法仅能发现人脸图像全局的欧式结构,无法发现隐藏在人脸图像中的非线性子流形结构,因此,对于姿态的变化,两种方法都无能为力,在识别率上性能基本相当。

随着主元数的继续增加,两种方法的识别错误率逐渐趋于一致。这可以通过主元对原始变量的贡献来解释。主元 Y_1, \dots, Y_m 与原始变量 X_1, \dots, X_p 的相关系数形成因子负荷阵 $A = (a_{ij})_{p \times m}$, 因子负荷阵的各列的平方和记为 $q_j^2 = \sum_{i=1}^p a_{ij}^2$, q_j^2 表示第 j 个主元对原始变量 X_1, \dots, X_p 的总影响。图7给出了 q_j^2 的趋势图,从图7可以看出,随着主元数的继续增加,两种方法的主元对原始变量的影响趋于一致,因此识别错误率趋于一致。

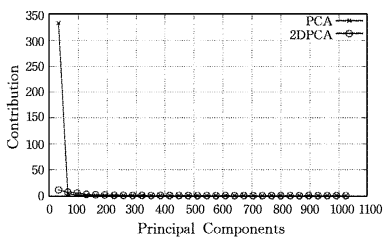


图7 q_j^2 的趋势图

4.3 表情变化实验

表情变化实验使用由卡梅隆大学机器人实验室构建的扩展 Cohn-Kanade (CK+) [12-13] 人脸表情库。该库包括 123 个人的 593 个人脸表情序列,共有 8 种基本表情(厌恶、高兴、惊讶、恐惧、愤怒、轻蔑、悲伤、正常),如图8所示。



图8 CK+人脸库中的8种基本表情

首先将人脸样本裁剪成为 32×32 大小,然后对其进行标准化,以期改善两种算法的性能。

本实验从库中选取正常表情(共 1051 张图像)作为训练集,分别采用生气表情子集(共 844 张图像)、高兴表情子集(共 1095 张图像)和惊奇表情子集(共 1117 张图像)作为测试集。识别采用最近邻决策规则,图9给出了两种方法在选取不同主元数时的实验结果。

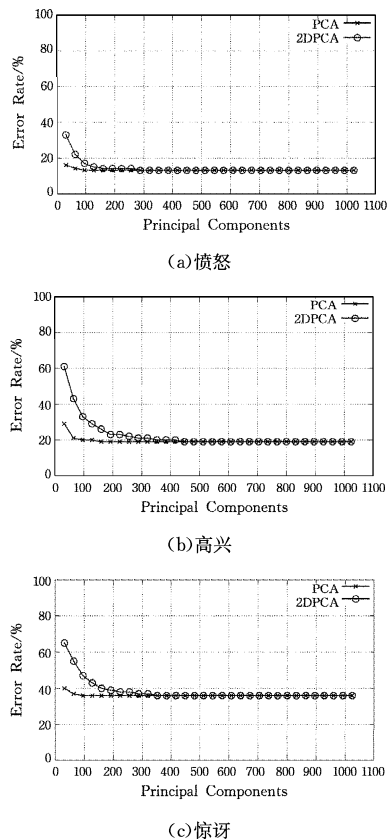


图9 两种方法在选取不同主元数时的识别错误率

在表情变化实验中,我们可以得到以下结论:从3个测试子集的错误率可以得到,随着主元数的增加,PCA与2DPCA的识别错误率趋于一致。其原因可以用两点来解释:

(1)选取两种方法的前 m 个主元作为投影轴,可以分别得到在不同 m 值下的重建图像(见图10)。从图10可以看到,随着主元数趋于最大,两种方法的重建图像质量趋于相同,这与图9中的变化趋势是相互对应的。

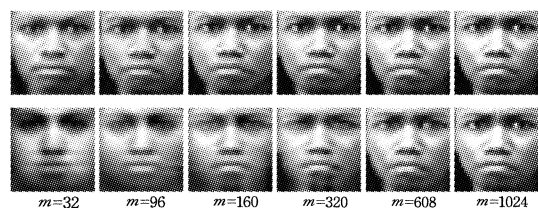


图10 不同主元数下的重建图像(第1行为PCA,第2行为2DPCA)

(下转第216页)

- [J]. Nature, 2004, 427(6970): 107-109.
- [3] 杜吉祥, 汪增福. 基于径向基概率神经网络的植物叶片自动识别方法[J]. 模式识别与人工智能, 2008, 21(2): 206-213.
- [4] LEE K, HONG K. An implementation of leaf recognition system using leaf vein and shape [J]. International Journal of Bio-Science and Bio-Technology, 2013, 214(2): 109-116.
- [5] 翟传敏, 汪青萍, 杜吉祥. 基于叶缘与叶脉分数维特征的植物叶识别方法研究[J]. 计算机科学, 2014, 5(2): 170-173.
- [6] CEVIKALP H, TRIGGS B. Face recognition based on image sets[C]// IEEE Conference on Computer Vision & Pattern Recognition. IEEE, 2010: 2567-2573.
- [7] Kim T K, Kittler J, CIPOLLA R. Discriminative Learning and Recognition of Image Set Classes Using Canonical Correlations [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2007, 29(6): 1005-1018.
- [8] WANG R, SHAN S, CHEN X, et al. Manifold-Manifold Distance with application to face recognition based on image set[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2008: 1-8.
- [9] WANG R, CHEN X. Manifold Discriminant Analysis[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2009: 429-436.
- [10] HARANDI M T, SANDERSON C, SHIRA ZI S, et al. Graph embedding discriminant analysis on Grassmannian manifolds for improved image set matching[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2011: 2705-2712.
- [11] HU Y, MIAN A S, OWENS R. Face Recognition Using Sparse Approximated Nearest Points between Image Sets [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2012, 34(10): 1992-2004.
- [12] DAVIS L S. Covariance discriminative learning: A natural and efficient approach to image set classification[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2012: 2496-2503.
- [13] YANG M, ZHU P, GOOL L V, et al. Face recognition based on regularized nearest points between image sets[C]// IEEE International Conference and Workshops on Automatic Face and Gesture Recognition. 2013: 1-7.
- [14] YAMAGUCHI O, FUKUI K, MAEDA K. Face Recognition Using Temporal Image Sequence[C]// Third IEEE International Conference on Automatic Face and Gesture Recognition. IEEE, 1998: 318-323.
- [15] ZHANG Y H, DU J X, WANG J, et al. Reverse Training for Leaf Image Set Classification[M]// Advanced Intelligent Computing Theories and Applications. Springer International Publishing, 2015: 233-242.

(上接第 206 页)

(2) 通常, 分类算法的性能由重建误差表征, 重建误差定义为 $\frac{1}{N} \sum_{i=1}^N \|A_i - \tilde{A}_i\|$, 在 PCA 中 $\tilde{A}_i = WY$, 在 2DPCA 中 $\tilde{A}_i = WY^T$. 图 11 给出了两种方法的重建误差. 随着主元数逐渐达到最大, 两种方法的重建误差趋于相等, 这与图 10 中的重建图像质量的趋势也是一致的.

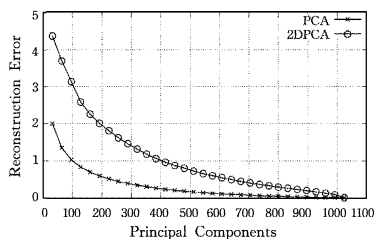


图 11 PCA 与 2DPCA 的重建误差

结束语 本文通过对 PCA 与 2DPCA 进行详细的理论分析, 证明了 PCA 与 2DPCA 在准则 $J(W)$ 下具有相同的最优目标值. 通过在 CMU-PIE 和 CK+ 库上进行了光照、姿态和表情变化的实验, 并对实验结果进行了理论分析, 最终证明了 PCA 与 2DPCA 在优化时具有相同的最优目标值, 即两种方法选取的主元数趋于最大时, PCA 的错误率与 2DPCA 趋于一致. 这有助于理解 PCA 和 2DPCA 之间的关系.

参考文献

- [1] PEARSON K. On Lines and Planes of Closest Fit to Systems of Points in Space [J]. Philosophical Magazine, 1901, 2(6): 559-572.
- [2] KIRBY M, SIROVICH L. Application of the KL Procedure for the Characterization of Human Faces [J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 1990, 12(1): 103-108.
- [3] TURK M, PENTLAND A. Eigenfaces for Recognition[J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71-86.
- [4] TURK M, PENTLAND A. Face recognition using eigenfaces [C]// Proc. IEEE Conf. on Computer Vision and Pattern Recognition. 1991: 586-591.
- [5] 甘玲, 邹宽中, 刘肖. 基于 PCA 降维的多特征级联的行人检测 [J]. 计算机科学, 2016, 43(6): 308-311.
- [6] YANG J, ZHANG D, FRANGI A F, et al. Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(1): 131-137.
- [7] 张焕龙, 郑卫东, 舒云星, 等. 基于区域协方差矩阵和 2DPCA 学习的视频跟踪方法研究 [J]. 计算机科学, 2015, 42(9): 278-281.
- [8] WANG L, WANG X, ZHANG X, et al. The equivalence of two-dimensional PCA to line-based PCA [J]. Pattern Recognition Letter, 2005, 26(1): 57-60.
- [9] GAO Q, ZHANG L, ZHANG D, et al. Comments on On Image Matrix Based Feature Extraction Algorithms [J]. IEEE Trans. on Syst. Man & Cybern. B Cybern., 2007, 37(5): 194-197.
- [10] 张贤达. 矩阵分析与应用[M]. 清华大学出版社, 2004: 54-57.
- [11] SIM T, BAKER S, BSAT M. The cmu pose, illumination, and expression (pie) database[C]// Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition. 2002.
- [12] KANADE T, COHN J F, TIAN Y L. Comprehensive database for facial expression analysis[C]// Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition. 2000: 46-53.
- [13] LUCEY P, CHON J F, KANADE T, et al. The extended cohn-kanade dataset (ck+): A complete expression dataset for action unit and emotion-specified expression[C]// Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis. 2010: 94-101.