

通用学习框架改进核 PCA 的单样本人脸识别

陈非¹ 黄山^{1,2} 张洪斌²

¹(四川大学电气信息学院 四川 成都 610065)

²(四川大学计算机学院 四川 成都 610065)

摘要 针对传统的人脸识别算法在每个人只有单个训练样本时识别性能严重下降的问题,提出了通用学习框架改进核主成分分析的单样本人脸识别算法。首先,选取一个合适的通用训练样本集,将各个单训练样本与通用训练样本集中某人的多训练样本按比例叠加;然后,利用经典的 KPCA 算法进行特征提取,将所有叠加后的训练样本和测试样本投影到特征子空间;最后,使用最近邻分类器完成最终的人脸识别。在 Yale 及 FERET 两大通用人脸数据库上的实验结果表明,相比其他几种较为先进的人脸识别算法,该算法取得了更好的单样本识别效果。

关键词 人脸识别 单样本每人 通用学习框架 最近邻分类器 核主成分分析

中图分类号 TP391 文献标识码 A DOI:10.3969/j.issn.1000-386x.2015.04.038

SINGLE SAMPLE FACE RECOGNITION WITH KERNEL PCA OPTIMISED BY GENERIC LEARNING FRAMEWORK

Chen Fei¹ Huang Shan^{1,2} Zhang Hongbin²

¹(School of Electrical Engineering and Information, Sichuan University, Chengdu 610065, Sichuan, China)

²(College of Computer Science, Sichuan University, Chengdu 610065, Sichuan, China)

Abstract For the problem that the recognition performance of traditional face recognition algorithms degrades seriously when each person has only one training sample, we propose a single sample face recognition algorithm which uses generic learning framework to improve kernel principle component analysis (KPCA). First, it selects a suitable generic training sample set and superposes each single training sample over the multiple training sample of a certain person in generic training set in proportion. Then, it uses typical KPCA to extract the features and projects all superposed training and testing samples onto feature subspace. At last, it uses the nearest neighbour classifier to complete the finale face recognition. Experiments results on two popular face databases of Yale and FERET show that the proposed algorithm achieves better recognition effect on single sample than several other relatively advanced face recognition algorithms.

Keywords Face recognition Single sample per person Generic learning framework Nearest neighbour classifier Kernel principle component analysis

0 引言

人脸识别技术已经在军事、安检、法律等领域得到了广泛的应用,目前已经存在许多人脸识别算法,如主成分分析(PCA)^[1,2]、独立成分分析(ICA)^[3]及线性判别分析(LDA)^[4,5]都可以成功地完成人脸识别。此外,无监督特征提取融合监督分类的算法也被引入到分类中,即核 PCA 加 LDA (KPCA + LDA)^[6]。大多数人脸识别算法在每个人有多个训练样本的情况下均能取得很好的识别效果,但是当每个人只有一个训练样本(即单训练样本)时,识别效果却不理想,因此,有效地解决单训练样本人脸识别问题成为了一大挑战^[7]。

针对单样本人脸识别问题,学者们提出了各种各样有效的识别算法,大致可以分为三类:无监督学习、虚拟样本扩张法、通用学习框架法^[8]。例如,文献[9]通过将每个单训练样本划分成若干个大小相等且互不重叠的局部小块,以每个类的各个

小块作为多训练样本,再利用 FLDA 进行特征提取,使得 FLDA 在单样本人脸识别问题中可用。后来,文献[10]提出了一种基于虚拟样本扩展的人脸识别算法,借助于相邻像素值差别很小的概念,将每个单训练样本抽取成多个低像素的人脸图像,每张人脸图像都与原图像非常相似,并利用传统的特征提取算法进行降维。文献[11]提出了一种基于 Candide-3 和支持向量机 SVM 的单训练样本人脸验证方法,该方法通过对重建模型的旋转产生姿态不同的数字人脸,将它们和原始样本一起作为训练数据,应用二分类问题的经典工具 SVM 进行分类匹配。文献[12]通过引用一个通用训练集学习判别特征,解决了单训练样本人脸识别中姿势变化的问题。上述各算法均在一定程度上解决了单训练样本人脸识别问题,然而,识别效果仍不理想。

受文献[11]“通过将单训练样本问题转化为多训练样本问

题求解”的启发,为了进一步改善单训练样本人脸识别性能,提出了一种通用学习框架改进 KPCA 的单样本人脸识别算法,然而,与文献[11]不同的是,本文算法引入了额外的通用训练样本集,并通过将单样本与通用训练样本集中每个人的多个样本按比例叠加,使得每个人都有多个训练样本,在基准数据库 Yale 及 FERET 的实验验证了所提方法的有效性和优越性。

1 算法设计与分析

本文算法的主要包括三个部分:通用学习框架运用、特征提取(即核主成分分析)、分类(识别),下面详细介绍各个部分。

1.1 通用学习框架

给定训练样本集 $A = \{A_i; i = 1, 2, \dots, m\}$, A_i 为 $M \times N$ 的人脸图像, m 为单训练样本的个数,通用人脸库 $C = \{C_{ij}; i = 1, 2, \dots, m; j = 1, 2, \dots, n\}$, C_{ij} 表示第 i 个人的第 j 个样本, n 表示 C 中每个人的样本数,首先将通用人脸库 C 中的各个人脸重构成大小为 $M \times N$ 的人脸,将 A 中的各个人脸与 C 中的各个人脸按照式(1)叠加后得到训练样本集 X 。

$$X_{ij}(x, y) = \alpha \times A_i(x, y) + (1 - \alpha) \times C_{ij}(x, y) \quad (1)$$

式中, α 表示原训练样本在训练样本集 X 中所占的比例,显然,如果 α 的值小于 0.5,则原样本在 X 中各个训练样本中所占的比例就小于通用样本,这样,训练集 B 中更多的代表的是通用样本集 C 中的样本信息,因此,我们规定, $0.5 \leq \alpha \leq 1$ 。 X_i 表示第 i 个人在训练样本集 X 中的第 j 个训练样本, X 中共有 $m \times n$ 个样本,即每个人都有 n 个训练样本。

如图 1 所示,第 1 张图像为 ORL 人脸库里的某个人的正面人脸图像,在我们的实验中作为单训练样本,第 2 至 9 张为 AR 人脸库中某人的 8 张表情、光照等条件各不相同的人脸图像,在我们的实验中为通用训练样本。



图 1 原训练样本

利用式(1), α 取 0.7, 将图 1 中的 9 张人脸图像变换得到了 8 张不同的人脸图像,如图 2 所示。



图 2 变换后的人脸图像

借助于通用学习框架,将单训练样本变成了多训练样本。

1.2 核主成分分析

特征提取是人脸识别系统的主要步骤,因为它能降低脸部图像的维度,本文算法使用核主成分分析(KPCA)提取高阶统计,算法过程描述如下:

第一步是集中给定的 n 个样本集 x_k , 其中, $x_k = [x_{k1}, x_{k2},$

$\dots, x_{kn}]^T \in R^n$ 。PCA 旨在找到最大化方差的投影 C , 即从协方差矩阵中找到特征值,问题变成求解式(2)的特征值:

$$\lambda v = C_x v \quad (2)$$

式中, $\lambda \geq 0$ 是特征值, $v \in R^n$ 是特征向量。

KPCA 中,输入空间的每个向量都通过非线性映射函数 $\varphi: R^n \rightarrow R^f$ 被投影到一个高维特征空间 R^f 中。

因此,PCA 的特征值可以计算如下:

$$\begin{aligned} \lambda v_\phi &= C_{\phi(x)} v_\phi = \left(\frac{1}{N} \sum_{i=1}^N \phi(x_i) \phi(x_j) \right) v_\phi \\ &= \frac{1}{N} \sum_{i=1}^N (\phi(x_i), v_\phi) \phi(x_j) \end{aligned} \quad (3)$$

v_ϕ 的所有 $\lambda \neq 0$ 的解都位于 $\phi(x_1), \dots, \phi(x_n)$ 范围内,存在系数 α_i 满足下列条件:

$$v_\phi = \sum_{i=1}^n \alpha_i \phi(x_i) \quad (4)$$

定义 $N \times N$ 核矩阵 K , 可表示为:

$$K_{i,j} = \phi(x_i) \cdot \phi(x_j) \quad (5)$$

将式(4)乘以 $\phi(x_j)$, 代入式(5), 得到:

$$N \lambda K \alpha = K^2 \alpha \quad (6)$$

该特征值问题的解可表示为以下更简单的特征值问题:

$$N \lambda \alpha = K \alpha \quad (7)$$

特征空间的投影由下式执行:

$$\begin{aligned} \phi(x)_{KPC}^k &= v_\phi^k \cdot \phi(x) = \sum_{i=1}^n \alpha_i^k (\phi(x_i) \phi(x_j)) \\ &= \sum_{i=1}^n \alpha_i^k K(x_i, x_j) \end{aligned} \quad (8)$$

因此,使用核函数不需要大量计算,就能提取 k 个非线性主成分。

1.3 识别

分类在任何人脸识别系统中都是非常关键的一步, k -最近邻分类器是最简单、使用最广泛的非参数化分类器,基于特征空间最接近训练实例对对象进行分类。特征提取完成后,本文算法计算测试样本特征与每个训练样本特征之间的欧氏距离,然后利用最近邻分类器^[13]完成最终的分类、识别。

综上,本文算法的流程如图 3 所示。

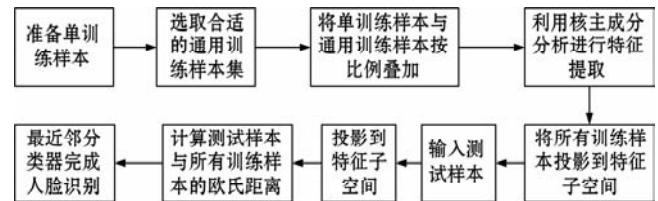


图 3 本文算法的流程

2 实验

所有的实验均在 4G 内存 Intel(R) Core(TM) 2.93 GHz, Windows XP 机器上完成,编程环境为 Matlab 7.0。

2.1 人脸数据库

Yale 人脸库^[14]包含了 15 个人的 165 张人脸,每人 11 张,包括了不同光照条件(灯光往左照射、往右照射、往中间照射),不同的面部表情(正常的、开心的、沮丧的、睡着的、惊讶的以及眨眼的),不同场景的(戴眼镜的和戴眼镜的),如图 4 所示为 Yale 人脸库中一个人的 11 副具有不同特征的人脸图像。



图4 Yale人脸库中某人的11副人脸图像

FERET人脸库^[15]包括ba、bb、bc、bd、be、bf、bg,共有200个人,每人7副人脸图像,它们是在不同的表情、视角及其光照强度下拍摄的。如图5所示,为其中一个人脸的7副图像,按照先行后列的顺序,对应为ba-bg中的一张人脸图像,选择ba作为训练样本,包括不同视角、不同表情和不同光照强度的bb至bg作为测试样本。



图5 FERET库中某人的7张人脸图像

2.2 实验结果

实验将原样本与通用训练样本均重新剪裁为 64×64 大小的图像,取Yale及FERET中每人的第一幅用于训练,剩下的用于测试,通用人脸库随机选择AR人脸库中的40个人,每人 n ($n=4,5,6,7,8,9,10$)张人脸图像,式(1)中的 α 取0.8,如表1所示分别为本文方法在Yale、FERET上随着通用训练样本数的变化取得的识别率。

表1 在Yale上取不同通用训练样本数所得识别率(%)

通用训练样本数 n	4	5	6	7	8	9	10
Yale	76.56	76.67	76.89	77.02	79.67	77.14	76.78
FERET-bb	33.5	35.0	34.5	34.5	37.5	33.5	34.0
FERET-bc	63.5	67.0	66.0	67.5	69.5	67.0	64.5
FERET-bd	71.0	70.5	71.5	71.5	71.0	71.0	71.5
FERET-be	41.0	41.5	41.0	42.0	44.0	43.0	42.5
FERET-bf	86.0	85.5	86.0	87.5	85.5	85.0	86.0
FERET-bg	32.0	31.5	30.5	31.0	33.5	31.5	30.5

从表1可以看到,随着通用训练样本数 n 的增大,Yale人脸库上的识别率呈升高的趋势,当 n 为8时,取得最高识别率79.67%,随着 n 继续增大,识别率呈下降的趋势。随着通用训练样本数 n 的变化,FERET人脸库上识别率的变化并不明显,bb至bg所得最高识别率分别为37.5%、69.5%、71.0%、44.0%、87.5%、33.5%。综合来看,随着 n 的变化,识别率的变化并不明显,由此可见本文算法的稳定性。

2.3 识别率比较

这部分将本文算法与几种较为先进算法的单训练样本识别率进行了比较,包括全局和局部特征的主成分分析(GLFP)^[2],增强Gabor特征的线性判别分析(EGFL)^[4],虚拟样本扩展的

2DPCA(VSE)^[10],分块Fisher线性判别分析(BFLD)^[9],以及通用Fisher线性判别分析(GFLD)^[12],即使用通用训练样本集的内、类间散布矩阵进行FLDA分类。

针对上面提到的几种比较算法,本文实验步骤简单介绍如下:GLFP算法中,取95%的能量来确定主成分数;EGFL算法中只有1个自由参数 α ,是人脸图像的投影组合权重,文献[4]中提到,当 α 的值在0.1至0.5之间时,EGFL的性能对 α 不敏感,因此,在实验中,取值0.3;BFLD算法中最重要的参数是分块数,采用了4种不同的分块方式(10×10 , 10×25 , 20×10 , 20×25),实验列表中选取了最佳的实验结果(10×25);类似的,在VSE中,人脸图像的分块对性能的影响很大,在实验中,采用了4种不同的分块数(16,32,40,72),选取了最佳的实验结果(72);GFLD算法中,先求出通用训练样本集的内、类间矩,然后利用FLDA进行分类,最近邻完成人脸的识别。几种算法在Yale及FERET上的实验结果如表2所示。

表2 各算法的单样本人脸识别率的比较(%)

算法	bb	bc	bd	be	bf	bg	平均	Yale
GLFP	31.5	61.5	65.5	29.5	80.5	6.0	46.58	70.17
EGFF	32.0	61.5	69.0	40.0	81.0	7.0	49.25	57.11
VSE	30.0	63.0	69.5	43.0	83.0	21.0	52.25	78.33
BFLD	27.5	62.5	69.0	41.0	80.5	4.5	48.00	71.83
GFLD	31.5	65.5	71.0	39.5	85.5	28.0	54.83	76.56
本文算法	37.5	69.5	71.0	44.0	87.5	33.5	58.17	79.67

从表2可以看出,在Yale及其FERET人脸库上,本文算法的识别率明显高于其它各个比较算法。其中,在Yale上,本文算法比GLFP、EGFL、BFLD算法的优势很明显,相比VSE算法,却仅高出了1.34个百分点,这不能体现出识别率的明显提高。但是,在FERET人脸库上,不论是看各个人脸库,还是看各个人脸库的平均识别率,本文算法都显得很优越,高出了经典的VSE算法6个百分点,甚至高出了GLFP算法近12个百分点。

Yale人脸库相对于FERET,光照、表情、场景等变化小得多,本文中引入的通用学习框架算法在Yale上的识别率提高不够显著,但是在FERET上却提高得很明显。由此可以看出,对处理光照、表情、场景等变化比较大的人脸库的时候,所提通用学习框架算法是很有效的。而在现实生活中,光照条件、人脸表情、场景等的变化是很大的,所以,通用学习框架算法的优点是明显的。

2.4 性能比较

为了更好地体现本文算法的优越性,将本文算法的复杂度与全局和局部特征的主成分分析(GLFP)^[2],增强Gabor特征的线性判别分析(EGFL)^[4],虚拟样本扩展的2DPCA(VSE-2DPCA)^[10],分块Fisher线性判别分析(BFLD)^[9],以及通用Fisher线性判别分析(GFLD)^[12]算法进行了对比,分别比较了训练时间复杂度、测试时间复杂度及空间复杂度,具体比较结果如表3所示,其中, m 和 n 分别表示图像矩阵的行数和列数, L 、 M 和 N 分别表示投影向量数、测试样本数和训练样本数。

从表3可以看出,与GLFP算法相比,本文算法在训练阶段的时间复杂度稍微高了一点,其它均相同;与EGFF算法相比,本文算法的各个复杂度均相当;与VSE-2DPCA算法相比,本文算法的训练阶段时间复杂度稍微高了点,但是测试阶段的时间复杂度比VSE-2DPCA算法低了一半;与BFLD、GFLD算法相比,本文算

法的训练阶段、测试阶段时间复杂度及总体空间复杂度均低了很多。在大大提高识别率的同时,本文算法仍然能够保持与其它相关算法相当甚至更优的复杂度,由此可见其优越性。

表3 各算法的复杂度比较

方法	训练阶段时间复杂度	测试阶段时间复杂度	总体空间复杂度
GLFP	$O(m^2 n^2 L)$	$O(mnL)$	$O(m^2 n^2)$
EGFF	$O(m^2 n^2 L + 2mnN^2)$	$O(mnL)$	$O(m^2 n^2)$
VSE-2DPCA	$O(m^2 n^2 L + mnN^2)$	$O(2mnL)$	$O(m^2 n^2)$
BFLD	$O(2m^2 n^2 L + mnN^2)$	$O(2mnL)$	$O(2m^2 n^2)$
GFLD	$O(2m^2 n^2 L + mnN^2)$	$O(2mnL)$	$O(2m^2 n^2)$
本文算法	$O(m^2 n^2 L + 2mnN^2)$	$O(mnL)$	$O(m^2 n^2)$

3 结 语

本文提出了一种基于通用学习框架改进核主成分分析的单样本人脸识别算法。通过引入通用训练样本,再利用经典的KPCA算法进行特征提取,最近邻分类器完成人脸分类,在Yale及FERET两大人脸数据库上的实验结果表明了本文算法的有效性及其可靠性。通用框架的引用,一定程度上提高了人脸识别率,但是找到一个非常合适的通用训练样本集,并且确定通用训练样本数是非常困难的事情,需要经过大量的实验才能确定。未来会将本文算法运用到其它数据集上,进一步提高单样本人脸识别率。

参 考 文 献

- [1] 于红芸,姜涛,关键. SAR 图像的快速核主成分分析识别方法[J]. 中国图象图形学报, 2012, 17(1): 137-141.
- [2] 苏煜,山世光,陈熙霖,等. 基于全局和局部特征集成的人脸识别[J]. 软件学报, 2010, 21(8): 1849-1862.
- [3] 柴智,刘正光. 应用复小波和独立成分分析的人脸识别[J]. 计算机应用, 2010, 30(7): 1863-1866.
- [4] 邹建法,王国胤,龚勋. 基于增强 Gabor 特征和直接分步线性判别分析的人脸识别[J]. 模式识别与人工智能, 2010 32(4): 477-482.
- [5] 尹洪涛,付平,沙学军. 基于 DCT 和线性判别分析的人脸识别[J]. 电子学报, 2009, 37(8): 2211-2214.
- [6] 卢桂馥,林忠,金忠. 基于核化图嵌入的最佳鉴别分析与人脸识别[J]. 软件学报, 2010, 22(7): 1561-1570.
- [7] 胡峰松,张茂军,邹北骥,等. 基于 HMM 的单样本可变光照,姿态人脸识别[J]. 计算机学报, 2009, 32(7): 1424-1433.
- [8] 杨军,高志升,袁红照,等. 基于 LBP 特征和贝叶斯模型的单样本人脸识别[J]. 光电子. 激光, 2011, 22(5): 763-765.
- [9] Chen S, Liu J, Zhou Z H. Making FLDA applicable to face recognition with one sample per person [J]. Pattern recognition, 2004, 37(7): 1553-1555.
- [10] 赵雅英,谭延琪,马小虎. 基于样本扩充和改进 2DPCA 的单样本人脸识别[J]. 计算机应用, 2011, 31(10): 2728-2730.
- [11] 马俊容. 单训练样本条件下人脸识别技术研究[D]. 湖南大学, 2009.
- [12] Hafiz F, Shafie A A, Mustafah Y M. Face Recognition From Single Sample Per Person by Learning of Generic Discriminant Vectors [J]. Procedia Engineering, 2012, 41(3): 465-472.
- [13] 杨军,高志升,袁红照,等. 基于 LBP 特征和贝叶斯模型的单样

本人脸识别[J]. 光电子. 激光, 2011, 22(5): 763-765.

- [14] 高涛,何明一. 改进投影梯度非负矩阵分解的单训练样本特征提取研究[J]. 电子与信息学报, 2010, 32(5): 1121-1125.
- [15] 王宪,慕鑫,张彦,等. 基于曲波域与核主成分分析的人脸识别[J]. 光电工程, 2011, 38(10): 98-102.

(上接第 136 页)

情况下,其分割结果中仍然出现大量误切,而在口语评测任务中对纯语音信号的误切是不能容忍的,且对之后的解码识别也将带来灾难性的后果,如表 4 所示本文算法分割所得音频的识别率要高近 15%。可见对于未知的突发性口语音频,基于固定阈值的音频分割算法其分割效果远不及本文算法。

4 结 语

本文通过分析中学口语测试音频的特征,并结合实时口语评测的特点,提出了一种新的基于自适应能量阈值与基频检测的音频分割算法,并给出了完整的音频分割流程。实验结果表明,该算法不仅能在确保考生有效口语音频完整的情况下,对实验音频进行快速准确的切分提高后续识别过程中的解码效率,并能去除音频噪声干扰,一定程度上提高音频的识别率,在实时口语评测应用中具有较强的适用性和较高的实用价值。然而,该算法也有它的局限性,该算法针对问答型口语任务,只是对单句的口语音频进行静音、噪音的切分,并未涉及到多句子段落音频切分后的句末判断,未实现多语句段落音频自动切分。总体来说,本文的音频切分算法达到了任务设计需求,取得了较好的分割效果。

参 考 文 献

- [1] Xie Shasha, Evanini K, Zechner K. Exploring content features for automated speech scoring [J]. Proceedings of the NAACL-HLT, Montreal, 2012:103-111.
- [2] Franco H, Neumeyer L, Yoon Kim, et al. Automatic Pronunciation Scoring for Language Instruction [C]. IEEE International Conference on Acoustics, Speech, and Signal Processing, 1997, 2: 1471-1474.
- [3] 张一彬,周杰,边肇祺,等. 一种新的基于分类的音频流分割方法[J]. 电子学报, 2006, 34(4): 612-616.
- [4] 张玉珍,夏肇霖,王建宇,等. 基于音频和文本融合的广告单元分割[J]. 南京理工大学学报:自然科学版, 2012, 36(3): 396-401.
- [5] 生龙,张洪斌. 二型模糊系统在音频信号分类中的应 [J]. 电子科技大学学报, 2013, 42(3): 436-441.
- [6] 王恩成,苏腾芳,袁开国,等. 哼唱检索中联合音高与能量的音符切分算法[J]. 计算机工程, 2012, 38(9): 4-7.
- [7] 郑玉婷,张文俊,韩彪. 一种适用于非特定哼唱方式的起音点检测算法[J]. 计算机工程, 2012, 38(10): 250-252.
- [8] 吕国云,蒋冬梅,蒋晓悦,等. 基于动态贝叶斯网络的音视频连续语音识别和音素切分[J]. 计算机应用, 2007, 27(7): 1670-1673.
- [9] 冯亚楠. 基于小波域 Teager 能量熵的音符起点检测算法研究[D]. 天津:天津大学电子信息工程学院, 2011.
- [10] 郑继明,张萍. 基于小波变换的音频分割[J]. 计算机工程与应用, 2011, 47(7): 139-142.
- [11] 刘秉权,徐师,李相前. 双阈值的特定英语音频句子边界检测[J]. 哈尔滨工业大学学报, 2010, 42(2): 259-263.
- [12] Rabiner L, Cheng M, Rosenberg A E, et al. A comparative performance study of several pitch detection algorithms [J]. IEEE Transaction on, Acoust, Speech, Signal Processing, 1976, ASSP-24(5): 399-417.