

基于差空间的双向 2DPCA 和 SVM 人脸识别算法

汪 洋¹ 严云洋^{1,2} 王洪元¹

(常州大学信息科学与工程学院 常州 213164)¹ (淮阴工学院计算机工程学院 淮安 223003)²

摘 要 为降低 SVM 人脸识别算法对样本进行训练和识别的时间,提出了一种改进的基于差空间的双向 2DPCA (Bidirectional two dimensions PCA)和 SVM 相结合的人脸识别算法。该方法充分考虑了表情和光照对人脸图像的影响,不但利用小波变换对人脸图像进行预处理,而且成功地把类内平均引入到双向 2DPCA 的计算中,并结合了 SVM 在分类识别方面的优势,有效节省了算法所需的时间。在 Yale 人脸库上的实验表明,它不但可以提高识别率,而且所用时间明显减少。

关键词 差空间,小波变换,双向 2DPCA,支持向量机,人脸识别

Face Recognition Based on Residual Space and SVM with Bidirectional Two Dimensions PCA

WANG Yang¹ YAN Yun-yang^{1,2} WANG Hong-yuan¹

(School of Information Science and Engineering, Changzhou University, Changzhou 213164, China)¹

(Faculty of Computer Engineering, Huaiyin Institute of Technology, Huai'an 223003, China)²

Abstract A novel face recognition algorithm was proposed to save the time of sample training and face recognition based on SVM. The new method is to recognize face based on residual space and SVM with Bidirectional two dimensions PCA. To avoid the influence of expression and light on face recognition, wavelet transform was used to process face images at first, then the within-class average was applied to the calculate two dimensions PCA. Furthermore SVM is in classification in order to effectively decrease the time of arithmetic. Experiments on Yale face data show that the new method can not only improve recognition rate, but also save the recognition time.

Keywords Residual space, Wavelet transform, Bidirectional two dimension PCA, SVM, Face recognition

1 引言

随着科技的发展,人脸识别这种生物特征识别以其方便、快捷、易于采集等特点受到人们的广泛关注,成为新时期下人工智能领域的研究热点^[1]。提取有效的鉴别特征和快速、准确的分类是决定识别效果的关键所在,通过对实际人脸识别系统的测试表明,姿态、表情和光照^[2]是影响人脸识别最重要的因素,人们一直致力于这方面的研究。Yang 等^[3]在前人工作的基础之上提出了 2DPCA 算法,该算法直接利用原始图像矩阵进行总体散度的计算,保留了人脸图像数据的固有结构,克服了传统 PCA 算法计算复杂的问题。在此基础上,李晓东、费树岷等^[4]引入了差空间的概念,较好改善了由于光照条件不同带来的不稳定性。二维 PCA 只在行方向上进行了 PCA 变换,而没有过多关注列方向上的变换,因此王心醉、李岩等^[5]提出了双向 PCA 算法。该方法在训练样本较多时取得了较好的识别效果,但对于小训练样本(训练样本数小于 5)的识别性能不是很理想。近年来,支持向量机(Support Vector Machines, SVM)这种新型的机器学习方法,以其能很

好地处理小样本问题,并且能工作在高维(甚至是无穷维)特征空间的独特优势受到人们的广泛关注。Thai Hoang Le 等^[6]将 2DPCA 与 SVM 相结合,在加快了特征提取速度的同时有效提高了识别率。但对于光照或表情变化明显的人脸图像则识别效果不是很理想。

基于以上的分析,本文结合了基于差空间的双向 2DPCA 人脸识别算法和 SVM 算法的特点,在充分考虑光照、姿态、表情等因数的影响下,利用小波变换处理图像具有对表情变化不太敏感的特点对图像进行预处理,克服了传统 SVM 人脸识别算法训练时间长的缺点。在 Yale 人脸库上的实验结果表明,该算法提高了识别性能,鲁棒性较好。

2 小波变换

为降低算法所需的时间,需要对训练样本图像进行压缩,以减少人脸图像的运算量并消除噪声。对人脸图像进行一次离散二维小波变换后,图像被分成 4 个子带:低通部分(LL),水平低通垂直高通部分(LH),水平高通垂直低通部分(HL),对角线方向上的高通部分(HH),如图 1 所示。其中 LL 包含

到稿日期:2012-02-28 返修日期:2012-05-20 本文受国家星火计划项目(2011GA690190),江苏省高校基金项目(08KJB520001,11KJD520003),江苏省“青蓝工程”,淮安市科技项目(HAG2010066, HAG2010030, HAG2011045, HAC201113)资助。

汪 洋(1985-),男,硕士生,主要研究领域为数字图像处理、模式识别, E-mail: wy12148@126.com; 严云洋(1967-),男,博士,教授,主要研究领域为数字图像处理、模式识别, E-mail: areyyk@163.com(通信作者); 王洪元(1960-),男,博士,教授,主要研究领域为模式识别与智能系统、数据融合。

了原始图像的主要特征, LH 和 HL 则反映了图像的细节信息, HH 则包含较多的噪声。研究发现, 小波分解后的低通以外的部分对人脸的最终鉴别也起着不可缺少的作用。



图1 人脸图像的小波变换

对一次离散小波分解的 4 个子带进行加权融合, 所取的权值组合分别为 I. 1/4, 1/4, 1/4, 1/4; II) 11/16, 3/16, 3/16, 1/16; III) 25/32, 3/32, 3/32, 1/32; IV) 57/64, 3/64, 3/64, 1/64; V) 1, 0, 0, 0。融合后的实验结果如图 2 所示。



图2 不同权值时融合后的图像

随着高频部分权值的减小, 图像越来越接近于低频时的图像。实验结果表明, 当高频部分权值较大时会引入噪声, 过小时会失去人脸的高频信息。实验发现, LL, LH, HL, HH 设置的权值分别为 25/32, 3/32, 3/32, 1/32 时融合的效果比较好。

3 算法原理

3.1 双向 2DPCA

3.1.1 双向 2DPCA 的基本原理

设 N 个训练样本为 $\{X_1, X_2, X_3, \dots, X_N\}$, 每个样本的维数是 $w \times h$, 第 i 个训练样本为 X_i 。将每个训练的样本图像矩阵看成是由 w 个 $1 \times h$ 的行向量构成的, 根据 2DPCA 的理论, 计算出训练样本图像行方向上的总体散度矩阵 G' :

$$G' = \frac{1}{Nw} \sum_{i=1}^N (X_i - \bar{X})^T (X_i - \bar{X}) \quad (1)$$

式中, \bar{X} 表示所有训练样本的均值矩阵。由矩阵 G' 的前 d 个最大特征值对应的特征向量 $(w_1, w_2, w_3, \dots, w_d)$ 组成最优投影矩阵, 令 $P_r = [w_1, w_2, w_3, \dots, w_d]$, 则 P_r 即为矩阵 G' 行方向的映射矩阵。

同理, 将每个训练的样本图像矩阵看成是由 h 个 $w \times 1$ 的列向量构成的。用类似的方法计算出训练样本列方向上的总体散度矩阵 G^c :

$$G^c = \frac{1}{Nh} \sum_{i=1}^N (X_i - \bar{X})(X_i - \bar{X})^T \quad (2)$$

取 G^c 的前 k 个最大特征值对应的特征向量组成映射矩阵 $P_c = [u_1, u_2, u_3, \dots, u_k]$, P_c 即为矩阵 G^c 列方向上的最优投影矩阵。

训练样本图像矩阵 X 的特征矩阵 Y 为:

$$Y = P_c^T X P_r \quad (3)$$

3.1.2 改进后的双向 2DPCA

鉴于人脸图像库中不同人之间的差别, 在总体散度矩阵的计算中引入类内平均。假设样本集 $\{X_1, X_2, X_3, \dots, X_N\}$ 共

有 L 类, N_1, N_2, \dots, N_L 分别表示每一类训练样本的数目。第 c 类训练样本图像可表示为 $\{X_1^c, X_2^c, \dots, X_{N_c}^c\}$, 其中 $X_i^c \in R^{w \times h}$, $i=1, 2, \dots, N_c, c=1, 2, \dots, L$ 。

第 c 类训练样本的类内平均矩阵 T_c 定义为:

$$T_c = \frac{1}{N_c} \sum_{i=1}^{N_c} X_i^c, c=1, 2, \dots, L \quad (4)$$

训练样本行方向上的总体散度矩阵 G' 为:

$$G' = \frac{1}{Nw} \sum_{c=1}^L \sum_{i=1}^{N_c} (X_i^c - T_c)^T (X_i^c - T_c) \quad (5)$$

同理, 训练样本图像列方向上的总体散度矩阵 G^c 为:

$$G^c = \frac{1}{Nh} \sum_{c=1}^L \sum_{i=1}^{N_c} (X_i^c - T_c)(X_i^c - T_c)^T \quad (6)$$

分别取 G' 的前 d 个最大特征值对应的特征向量组成训练样本行方向上的最优映射矩阵 $P_r' = [w_1', w_2', w_3', \dots, w_d']$ 和 G^c 的前 k 个最大特征值对应的特征向量组成训练样本列方向上的最优映射矩阵 $P_c' = [u_1', u_2', u_3', \dots, u_k']$, 则训练样本图像矩阵 X 的特征矩阵 Y' 为:

$$Y' = P_c'^T X P_r' \quad (7)$$

3.2 差空间中的差图像

为尽可能降低人脸图片中的光照信息对人脸识别性能的影响, 构造了差图像, 具体步骤如下:

(1) 用改进的双向 2DPCA 的方法计算出训练样本的特征矩阵 Y' 。

(2) 结合投影后的特征矩阵 Y' , 对原始训练样本 X 进行重构, 设重构后的样本矩阵为 X' , 则:

$$X' = P_r'^T Y' P_c' \quad (8)$$

通过观察重构后的样本图像不难发现, 除人脸固有的内在信息外, 还有丰富的光照信息, 如何去掉这些非人脸的本质信息, 将是提高人脸识别率的关键所在。

(3) 为去掉光照信息, 定义了差图像 Q :

$$Q = X - X' \quad (9)$$

实验表明, 差空间中的每幅差图像的确保留了人脸内在固有信息, 且摒弃了大量的光照信息。

使用 MATLAB2009b 在 Yale 人脸库上进行了实验, 结果如图 3 所示。

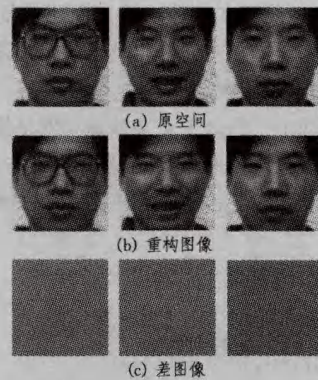


图3 差图像构造

4 支持向量机

4.1 支持向量机的基本原理

支持向量机 (Support Vector Machine, SVM) 是 Cortes 和 Vpanik 于 1995 年首先提出的, 它建立在统计学习理论的 VC

维理论和结构风险最小原理基础上,根据有限的样本信息在模型的复杂性和学习能力之间寻求最佳折衷,以期获得最好的称泛化能力^[7]。

SVM最初是用来对二分类问题进行分类的,在线性可分的情况下,对于训练集 (x_i, y_i) ,其中 $x_i \in R_N, y_i \in \{-1, 1\}, i=1, 2, 3, \dots, n$,目标是找到一个超平面能够将两类样本完全分开且两类间的间隔最大。不妨设超平面的方程为: $w \cdot x + b = 0$,将 $w \cdot x + b > 0$ 的归为1类,而 $w \cdot x + b < 0$ 的则归为-1类。两类间隔的最大化等价于最小化:

$$J(w) = \frac{1}{2} \|w\|^2 \quad (10)$$

约束条件:

$$y_i(w \cdot x_i + b) \geq 1, \forall i \in \{1, 2, \dots, n\}$$

引入 Lagrange 乘子 α_i , 得到式(1)的 Wolfe 对偶形式为:

$$L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (11)$$

在解得 α 之后,平面参数 w 和 b 便可以由对偶问题的解 α 来确定。由此可以得到支持向量机的最优分类函数:

$$f(x) = \text{sgn}(\sum_{i=1}^n \alpha_i y_i (x_i \cdot x) + b) \quad (12)$$

对于非线性可分的情况,则需要利用核函数将原样本映射到高维空间中转化为高维空间中的线性可分问题,其对应的目标函数变为:

$$J(w, b, \epsilon) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \epsilon_i \quad (13)$$

式中, C 为惩罚因子, ϵ 是松弛变量, 此时约束条件变为:

$$y_i(w \cdot x_i + b) \geq 1 - \epsilon_i, \epsilon_i \geq 0, \forall i \in \{1, 2, \dots, n\}$$

而最优的分类函数则变为:

$$f(x) = \text{sgn}(\sum_{i=1}^n \alpha_i y_i K(x_i \cdot x) + b) \quad (14)$$

式中, $K(x_i \cdot x)$ 为核函数。常见的核函数有:

(1) 多项式核函数: $K(x, x') = (x \cdot x' + c)^d$;

(2) 径向基核函数: $K(x, x') = \exp(-\|x - x'\|^2 / \sigma^2)$;

(3) Sigmoid 核函数: $K(x, x') = \tanh(k(x, x') + \nu)$ 。

4.2 核函数的选取以及参数的确立

多项式核函数虽然计算简单,但是需要计算内积,由此可能会产生溢出之类的计算问题; Sigmoid 核函数的行为类似于径向基核函数,但是其在参数选择方面比较复杂,有 k 和 ν 两个参数需要确定。本文在充分比较了各个核函数的优缺点之后,选用了径向基核函数。

借助林智仁博士^[8]的基于交叉验证和网格搜索参数选择方法来选取最优参数。在 Faruto 等人^[9]的工作基础之上,选用 Lib svm_mat_2.89 进行了实验,实验结果如图 4 所示,其中, $g=1/\sigma^2$ 。

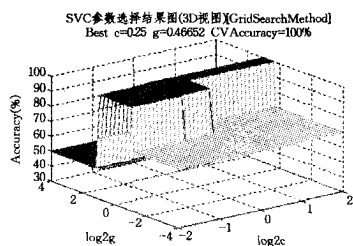


图 4 最优参数的选取

4.3 多类分类器的构造

人脸图像库是一个多分类问题,有别于二分类,需要将二

分 SVM 进行推广,本文采用的是一对多最大响应策略。

5 算法的流程

训练的流程:(1)通过双向 2DPCA 对训练样本降维。(2)利用 3.2 节中的方法构造差图像,对所得差图像再次降维,得到最大特征值对应的特征向量组成的特征矩阵。(3)将训练图像投影在特征矩阵上得到的训练样本的特征放入 SVM 训练器中进行训练。识别的流程:(1)通过投影的方法得到待识别图像的特征。(2)利用 SVM 分类器进行分类识别。整个算法流程图如图 5 所示。

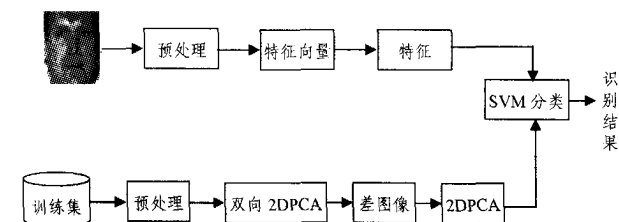


图 5 算法流程图

6 实验结果及分析

实验选用 Yale 人脸库,它包含 15 人,每人 11 幅图片。Yale 库中人的脸部表情丰富,细节变化明显,有笑、微笑、不笑;眼睛有睁有闭;还有戴或不戴眼镜之分;另外,人脸的光照也有相当程度的变化,人脸的尺度也略有变化。图 6 是 Yale 库中某人的 11 幅图像。



图 6 Yale 库中的人脸样本

在每人 11 幅图像中分别以 1~9 幅图像作为训练样本,其余作为测试样本进行实验,并与文献^[10]中的方法在识别率和所用时间上进行比较,实验所得的数据如表 1 和表 2 所列。

表 1 训练样本数对识别率的影响

| 训练样本数 | 识别率(%) | | | | | | | | |
|--------------|--------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Wave+PCA+SVM | 65.83 | 81.48 | 83.81 | 92.38 | 91.11 | 88.00 | 91.67 | 86.67 | 90.00 |
| 本文方法 | 84.33 | 88.89 | 91.12 | 92.38 | 94.44 | 95.00 | 93.33 | 91.11 | 90.00 |

表 2 训练样本数对时间的影响

| 算法 | 算法所用时间(s) | | | | | | | | |
|--------------|-----------|--------|--------|--------|--------|--------|--------|--------|--------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Wave+PCA+SVM | 9.844 | 11.672 | 12.625 | 17.907 | 20.11 | 28.36 | 26.687 | 26.39 | 19.86 |
| 本文方法 | 9.625 | 10.64 | 11.656 | 16.422 | 15.985 | 20.562 | 22.359 | 18.516 | 12.985 |

实验结果表明,在训练样本小于 5 时,本文算法识别率较好,与融合小波的 PCA+SVM 人脸识别算法相比,在时间和识别率上都有明显的提高。

在训练样本为每人 5 幅时,本文方法在 Yale 人脸库上与一些经典方法在样本训练时间和测试时间以及识别率上的比较如表 3 所列。

表 3 算法的比较

| 选用方法 | 训练时间(s) | 识别时间(s) | 识别率(%) |
|--------------|---------|---------|--------|
| PCA+SVM | 82.607 | 19.377 | 90.00 |
| Wave+PCA+SVM | 16.490 | 3.619 | 91.11 |
| 2DPCA+SVM | 13.665 | 5.582 | 93.33 |
| 本文方法 | 12.948 | 3.037 | 94.44 |

从以上比较的数据上来看,本文的算法由于采用了小波变换和双向 2DPCA 的方法,因此在同等条件下,算法所需的识别时间比传统的 SVM 方法要低得多,有效解决了 SVM 算法耗时较多的问题。

综合以上实验结果可以看出,本文方法简单、稳定,识别速度快,识别率也较高。

结束语 为进一步提高人脸识别的性能,本文提出了一种改进的基于差空间的双向 2DPCA+SVM 人脸识别算法,其利用加权小波变换对图像进行压缩,并引入差空间和类内平均的概念,从而有效缩短了算法所需的时间,提高了识别率,算法的鲁棒性较好。值得注意的是,本文是在静态人脸库上进行的实验,当人脸处于运动状态时,算法的有效性将有待进一步研究。

参 考 文 献

[1] He Xiao-fei, S Y, He Yu-xiao, et al. Face Recognition Using

(上接第 260 页)

[4] Deutsch G K, Dougherty R F, Bammer R, et al. Children's reading performance is correlated with white matter structure measured by diffusion tensor imaging[J]. *Cortex*, 2005, 41(3): 354-363

[5] Jonasson L, Hagmann P, Bresson X, et al. White matter mapping in DT-MRI using geometric flows[C]//Proc. of Intl' Conf. on EUROCAST 2003, LNCS 2809, 2003: 585-596

[6] Wiegell M R, Tuch D S, Wedeen V J, et al. Automatic segmentation of thalamic nuclei from diffusion tensor magnetic resonance imaging[J]. *NeuroImage*, 2003, 19(2): 391-401

[7] Wang Z Z, Vemuri B C. Tensor field segmentation using region based active contour model[C]//Proc. of Intl' Conf. on ECCV 2004, LNCS 3024, 2004: 304-315

[8] Lenglet C, Rousson M, Deriche R. DTI segmentation by statistical surface evolution[J]. *IEEE Transactions on Medical Imaging*, 2006, 25(6): 685-700

[9] Yushkevich P A, Piven J, Hazlett H C, et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability[J]. *NeuroImage*, 2006, 31(3): 1116-1128

[10] Awate S P, Zhang H, Gee J C. A fuzzy nonparametric segmentation framework for DTI and MRI analysis; with application to

Laplacianfaces [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(3): 328-340

[2] Tan Xiao-yang, Enhanced B T. Local Texture Feature Sets for Face Recognition under Difficult Lighting Condition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 19(6): 1653-1650

[3] Yang Jian, Zhang D. Senior Member 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

[4] 李晓东, 费树岷, 张涛. 基于差空间和最大散度差鉴别分析的人脸识别方法[J]. *东南大学学报*, 2009, 39(6): 1130-1135

[5] 王心醉, 李岩, 郭立红, 等. 基于双向 PCA 和 K 近邻的人脸识别算法[J]. *解放军理工大学学报*, 2010, 11(6): 623-627

[6] Le T H, L B. Face Recognition based on SVM and 2DPCA[J]. *International Journal of Signal Processing Image Processing and Pattern Recognition*, 2011, 4(3): 85-93

[7] 张静, 王艳平, 薛桂香. 数字图像处理与计算机视觉[M]. 北京: 人民邮电出版社, 2010: 511-544

[8] Chang C-C, Lin C-J. LIBSVM: a library for support vector machine[OL]. <http://www.csie.ntu.edu/~cjlin/libsvm>, 2001

[9] Faruto. SVM 视频讲解汇总[OL]. <http://www.ilovematlab.cn/thread-62253-1-1.html>, 2010

[10] 聂会星, 梁坤, 徐枞巍. 基于小波变换和支持向量机的人脸识别研究[J]. *合肥工业大学学报*, 2011, 34(2): 208-211

DTI-tract extraction[J]. *IEEE Transactions on Medical Imaging*, 2007, 26(11): 1525-1536

[11] 吴占雄, 朱善安, 贺斌. 扩散张量成像中脑胼胝体结构图像的分割算法[J]. *浙江大学学报*, 2011, 45(1): 163-167

[12] Zhang K H, Zhang L, Song H H, et al. Active contours with selective local or global segmentation; a new formulation and level set method[J]. *Image and Vision Computing*, 2010, 28(4): 668-676

[13] Tomán H, Tornai R, Zichar M. Complex fiber visualization[J]. *Annales Mathematicae et Informaticae*, 2007, 34: 103-109

[14] Partridge S C, Mukherjee P, Henry R G, et al. Diffusion tensor imaging; serial quantitation of white matter tract maturity in premature newborns[J]. *NeuroImage*, 2004, 22(3): 1302-1314

[15] Huang H, Zhang J Y, Jiang H Y, et al. DTI tractography based parcellation of white matter: Application to the mid-sagittal morphology of corpus callosum[J]. *NeuroImage*, 2005, 26(1): 195-205

[16] Chan T F, Vese L A. Active contour without edges[J]. *IEEE Transactions on Image Processing*, 2001, 10(2): 266-277

[17] de Luis-García R, Alberola-López C. Mixtures of gaussians on tensor fields for DT-MRI segmentation [C]//Proc. of Intl' Conf. on MICCAI 2007, LNCS 4791, 2007: 319-326