ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.17, No.12, December 2006, pp.2508−2517
DOI: 10.1360/jos172508
© 2006 by *Journal of Software*. All rights reserved.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

基于局部 Gabor 变化直方图序列的人脸描述与识别*

张文超[1+], 山世光[2], 张洪明[1], 陈 杰[1], 陈熙霖[2], 高 文[1,2]

[1](哈尔滨工业大学 计算机科学与技术学院,黑龙江 哈尔滨 150001)

[2](中国科学院 计算技术研究所,北京 100080)

Histogram Sequence of Local Gabor Binary Pattern for Face Description and Identification

ZHANG Wen-Chao[1+], SHAN Shi-Guang[2], ZHANG Hong-Ming[1], CHEN Jie[1], CHEN Xi-Lin[2], GAO Wen[1,2]

[1](School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

[2](Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100080, China)

+ Corresponding author: Phn: +86-10-58858300 ext 316, Fax: +86-10-58858301, E-mail: wczhang@jdl.ac.cn, http://www.jdl.ac.cn

Zhang WC, Shan SG, Zhang HM, Chen J, Chen XL, Gao W. Histogram sequence of local Gabor binary pattern for face description and identification. *Journal of Software*, 2006,17(12):2508−2517. http://www.jos.org.cn/1000-9825/17/2508.htm

Abstract: In this paper, a method for face description and recognition is proposed, which extracts the histogram sequence of local Gabor binary patterns (HSLGBP) from the magnitudes of Gabor coefficients. Since Gabor feature is robust to illumination and expression variations and has been successfully used in face recognition area. First, the proposed method decomposes the normalized face image by convolving the face image with multi-scale and multi-orientation Gabor filters to extract their corresponding Gabor magnitude maps (GMMs). Then, the local binary patterns (LBP) operates on each GMM to extract the local neighbor pattern. Finally, the input face image is described by the histogram sequence extracted from all these region patterns. The proposed method is robust to illumination, expression and misalignment by combing the Gabor transform, LBP and spatial histogram. In addition, this face modeling method does not need the training set for statistic learning, thus it avoids the generalizability problem. Moreover, how to combine the statistic method in the stage of classification and propose statistic Fisher weight HSLGBP matching method are discussed. The results compared with the published results on FERET face database of changing illumination, expression and aging verify the validity of the proposed method.
Key words: face recognition; Gabor filter; local binary pattern (LBP); histogram

摘 要: 提出了一种在 Gabor 变换幅值域内提取局部变化模式空间直方图序列(histogram sequence of local Gabor binary patterns,简称 HSLGBP)的人脸描述及其识别方法.鉴于 Gabor 特征对光照、表情等变化比较鲁棒,并已在人脸识别领域得到成功应用,首先对归一化的人脸图像进行多方向、多分辨率 Gabor 小波滤波,并提取其对应不同方向、不同尺度的多个 Gabor 幅值域图谱(Gabor magnitude map,简称 GMM),然后在每个 GMM 上采用

* Supported by the National Natural Science Foundation of China under Grant Nos.60332010, 60673091 (国家自然科学基金); the "100 Talents Program" of CAS (中国科学院"百人计划"); the ISVISION Technologies Co., Ltd (上海银晨智能识别科技有限公司资助项目)

Received 2005-07-13; Accepted 2005-12-31

局部二值模式(local binary pattern,简称 LBP)算子抽取局部邻域关系模式,最后由这些模式的区域直方图形成的序列来描述人脸.Gabor 变换、LBP、空间区域直方图的采用使得该方法对光照变化、表情变化、误配准等具有良好的鲁棒性.而且,这种人脸建模方法不需要基于训练集合进行统计学习,因而不存在推广性问题.同时,进一步探讨了如何在分类器设计阶段与统计方法进行结合的问题,提出了统计 Fisher 加权的 HSLGBP 匹配方法.在通过 FERET 人脸库光照、表情和时间变化测试集上与已发表的实验结果进行对比,充分验证了该方法的有效性.

关键词: 人脸识别;Gabor 滤波器;局部二值模式(LBP);直方图
中图法分类号: TP391 文献标识码: A

作为模式识别、图像分析与理解等领域的典型研究课题,人脸识别不仅在理论上具有重要价值,而且在安全、金融等领域具有重要的应用前景,因此在学术界和产业界都受到了广泛的关注,目前已经出现了一些实用的商业系统.然而,由于图像采集条件和人脸自身属性的变化,例如图像采集时的光照、视角、摄像设备的变化,人脸的表情、化妆、年龄变化等,都可能使得同一人的不同照片表观差别很大,造成识别上的困难.因此,提高人脸识别系统对这些变化的鲁棒性成为该领域研究人员的重要目标之一[1,2].为实现鲁棒的识别,可以从人脸建模、分类器设计等不同角度入手:前者试图从寻求对各种外界条件导致的图像变化具有不变性的人脸描述入手;而后者则试图使得分类器对特征的变异有足够的鲁棒性.另外,也可以同时从两个角度入手解决问题.

在人脸识别的早期,基于表观(appearance)的人脸识别方法往往直接采用图像灰度作为特征进行分类或者特征选择的基础,如 Correlation[3],Eigenfaces[4],Fisherfaces[5]等.而近年来,对图像灰度进行多尺度、多方向的 Gabor 小波变换逐渐成为主流的思路之一,这主要是因为 Gabor 小波可以很好地模拟大脑皮层中单细胞感受野的轮廓,捕捉突出的视觉属性,例如空间定位、方向选择等[6].特别是 Gabor 小波可以提取图像特定区域内的多尺度、多方向空间频率特征,像显微镜一样放大灰度的变化,这样,人脸图像中的眼睛、鼻子和嘴以及其他一些局部特征被放大.因此,采用 Gabor 变换来处理人脸图像,可以增强一些关键特征,以区分不同人脸图像.Gabor 小波也因此在人脸识别中得到了广泛的应用,如弹性图匹配[7]、基于 Gabor 特征的 Fisher 判别[8]、基于 AdaBoost 的 Gabor 特征分类[9]等.但是,由于多尺度、多方向的 Gabor 分解使得数据的维数大量增加,尤其当图像尺寸偏大时更为严重,为避免维数灾难问题必须进行降维.弹性图匹配方法仅对人脸图像中部分关键特征点进行 Gabor 变换,并将人脸描述为以这些特征点位置为顶点、以其 Gabor 变换系数为顶点属性、以其关键点位置关系为边属性的属性图,从而将人脸识别问题转化为图匹配问题[7],但该方法对特征点的选择和配准有较高的要求.直接降维的方法是在 Gabor 变换系数下采样,然后采用 Fisher 判别分析方法进一步提取特征[8],但简单下采样很可能造成一些重要信息的丢失.针对该问题,文献[9]提出了一种基于 AdaBoost 对高维 Gabor 特征进行选择的降维方法,采用机器学习的方法更加客观地利用 Gabor 特征,从而能在有效降维的同时提高识别性能.

上述方法在一些人脸库上取得了很好的识别结果,但由于它们采用了统计或者学习的策略,因此都不可避免地遇到推广性的问题,即算法性能在很大程度上依赖于测试集和训练集之间数据分布的相似程度.尽管统计学习理论以该问题为核心进行了深入的理论探讨[10],但在模式识别算法应用实践上,该问题仍然很棘手,尤其是对很多人脸识别实用系统而言,往往不可能获得待识别人的多个样本,这就意味着不可能对它们进行有效的训练.例如,在护照、驾照、身份证等的验证、大规模人脸图像库上的照片检索等应用中,每个人都只提供了单幅图像,很难进行针对性的训练.解决该问题的可能思路有两条:其一是仍然采用统计、学习策略,寄希望于其泛化能力;其二则是考虑非统计学习的策略,转向采用直接的模型匹配方法.近年来,基于局部二值模式(local binary pattern,简称 LBP)[11]的人脸识别方法受到人们的关注,该方法来源于纹理分析领域.它首先计算图像中每个像素与其局部邻域点在亮度上的序关系,然后对二值序关系进行编码形成局部二值模式,最后采用多区域直方图作为图像的特征描述.该方法在 FERET 人脸图像数据库上取得了很好的识别性能.

LBP 方法本质上提取的是图像边缘、角点等局部变化特征,它们对于区分不同的人脸是很重要的.但是,边缘特征有方向性和尺度差异,角点特征也同样有不同的尺度,或者说边缘、角点等图像特征的方向性和尺度包

含了更多的可以区分不同人脸的细节信息,而原始的 LBP 算子却不能提取这些特征.基于此,本文提出了一种称为局部 Gabor 二值模式直方图序列的人脸描述及其相应的识别方法(histogram sequence of local Gabor binary patterns,简称 HSLGBP).该方法首先利用 Gabor 变换提取多方向、多尺度的局部图像特征,然后应用 LBP 算子对这些特征进行编码,最后采用空间直方图进行人脸建模,人脸的分类识别则通过直方图匹配来实现.在此基础上,本文还进一步考察了基于 Fisher 准则对不同区域加权的直方图匹配策略,以进一步提高识别精度.与直接从图像的灰度计算 LBP 空间直方图相比,Gabor 变换和 LBP 的结合提取了更多方向、更多尺度的局部特征,从而有效地增强了空间直方图的表示能力,我们在 FERET 等人脸库上的实验有力地表明了这一点.尤其是使用加权 HSLGBP 在 FERET 人脸库上的表情、光照和时间变化几个测试集上均取得了目前已知的最好结果.这进一步表明 HSLGBP 对人脸图像条件的变化是鲁棒的,而且具有很好的判别能力.

1 基于 HSLGBP 的人脸描述方法

本文提出的 HSLGBP 人脸描述提取过程如图 1 所示,主要包括以下步骤:(1) 根据手工标定人眼位置对输入人脸图像归一化;(2) 将 40 个不同尺度、不同方向的 Gabor 滤波器分别与归一化图像卷积,取每个卷积结果的幅值部分,这里称为 Gabor 幅值图谱(Gabor magnitude map,简称 GMM);(3) 对每个 GMM 应用 LBP 运算得到局部 Gabor 二值模式(local Gabor binary pattern,简称 LGBP)图谱;(4) 将每个 LGBP 图谱划分为互不重叠且具有特定大小的多个矩形区域,对每个区域计算直方图;(5) 将 GMM 的所有区域的直方图串接为一个直方图序列,作为人脸图像的描述.下面详细阐述这一过程.

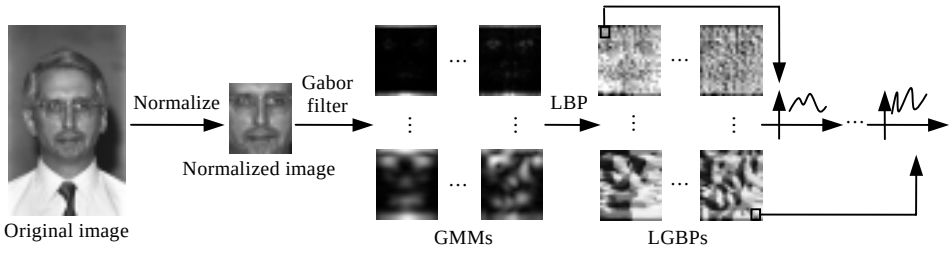


Fig.1 The framework to extract the proposed HSLGBP face representation

图 1 本文提出的 HSLGBP 人脸描述提取方法框架

1.1 Gabor幅值图谱

考虑到如前所述 Gabor 变换的优良特性,我们使用 Gabor 滤波器来分解输入的人脸图像.本文使用的 Gabor 滤波器如式(1)所示[12].

$$\psi_{\mu,\nu}(z)=\frac{\left\|k_{\mu,\nu}\right\|^2}{\sigma^2}\mathrm{e}^{\left(-\left\|k_{\mu,\nu}\right\|^2\|z\|^2/2\sigma^2\right)}\left[\mathrm{e}^{ik_{\mu,\nu}z}-\mathrm{e}^{-\sigma^2/2}\right] \tag{1}$$

其中,$z=(x,y)$;$\|\cdot\|$表示范数运算;$k_{\mu,\nu}=k_\nu \mathrm{e}^{i\phi_\mu}$,$k_\nu=k_{\max}/f^\nu$,$\phi_\mu=\pi\mu/8$,$\mu$和 ν 分别表示 Gabor 滤波器的方向和尺度.

人脸图像的 Gabor 特征由人脸图像和 Gabor 滤波器的卷积得到.令 $f(x,y)$表示人脸图像的灰度分布,那么 $f(x,y)$和 Gabor 滤波器的卷积可定义为

$$G(x,y,\nu,\mu)=f(x,y)*\psi_{\nu,\mu}(z) \tag{2}$$

其中,'*'表示卷积运算.这样,由一系列的 $\nu\in\{0,1,...,4\}$ 和 $\mu\in\{0,1,\ldots,7\}$即可得到人脸图像的多层 Gabor 滤波器分解表示.

1.2 局部Gabor二值模式

Gabor 变换得到的是由实部和虚部组成的复数,包含幅值和相位谱.其中,相位谱随着空间位置呈周期性变化[7],因而通常认为不适合作为人脸特征.而幅值的变化相对平滑而稳定,因此,本文仅用变换后的幅值作为人脸

特征的描述.由于与 40 个 Gabor 滤波器进行卷积,Gabor 特征的维数相对于原始图像维数急剧增加,必须进行降维.常用的降维方法有两种:一种是首先进行均匀下采样,然后进行特征提取(如主成分分析)[8],但下采样可能会丢失一些重要的判别特征;另一种方法,如弹性图匹配[7],只对人脸图像上选择的一些特征点做 Gabor 变换,但这样既对特征点的定位要求非常高,又对特征点的选择敏感.本文采用 LBP 算子来编码 Gabor 幅值的邻域变化,并用直方图分析 LBP 编码后的局部变化属性,这样直接对 GMM 分析,避免了因下采样造成的信息损失,同时又通过直方图统计达到了降维的目的.

近年来,LBP 算子在人脸识别领域中受到关注[11,13,14],该算子[15]对图像每个像素 f_c 的 8 邻域采样,每个采样点 $f_p(p=0,1,\ldots,7)$与中心像素 f_c 做灰度值二值化运算 $S(f_p-f_c)$.

$$S(f_p-f_c)=\begin{cases}1, & f_p\ge f_c\\ 0, & f_p<f_c\end{cases} \tag{3}$$

其中 f_c 表示该中心像素的灰度值;f_p 表示采样点的灰度值.然后,通过对每个采样点赋予不同的权系数 2^p 来计算该 f_c 的 LBP 值,

$$LBP=\sum_{p=0}^{7}S(f_p-f_c)2^p \tag{4}$$

LBP 运算刻画了局部图像纹理的空间结构.本文对 Gabor 幅值进行 LBP 运算,简记为 LGBP.

$$LGBP=\sum_{p=0}^{7}S(G_p(x,y,\mu,\nu)-G_c(x,y,\mu,\nu))2^p \tag{5}$$

1.3 LGBP空间区域直方图序列(HSLGBP)

人脸表观由于表情、时间等变化呈现多样性,带来识别上的困难.基于全局的人脸表示方法对这种面部的局部变化不鲁棒,而基于区域的分析通常会较好地解决这一问题[16].基于此,我们将人脸图像划分为多个区域进行分析,即将 LGBP 划分为多个不相交的矩形区域,并用直方图刻画每个区域的分布属性.这样,高维 LGBP 变为低维的直方图,人脸图像用所有直方图串接而成的直方图序列描述.

我们采用直方图对 LGBP 进行统计,同时通过选择合适的直方图 Bin 数达到对 LGBP 降维的目的.灰度范围是[0,L−1]的图像 $f(x,y)$的直方图可定义为

$$\boldsymbol{h}_i=\sum_{x,y}I\{f(x,y)=i\},\ i=0,1,...,L-1 \tag{6}$$

其中 i 表示第 i 个灰度级;$\boldsymbol{h}_i$ 是具有第 i 级灰度的像素的数目,并且

$$I\{A\}=\begin{cases}1, & A\text{ is true}\\ 0, & A\text{ is false}\end{cases} \tag{7}$$

将每个 LGBP 图谱划分为 m 个区域 $R_0,R_1,\ldots,R_{m-1}$,根据式(8)从每个区域提取直方图:

$$H_{\nu,\mu,R_j,i}=\sum_{(x,y)\in R_j}I\{LGBP(x,y,\mu,\nu)\}I\{(x,y)\in R_j\},\ i=0,1,...,L-1;\ j=0,1,...,m-1;\ \nu=0,1,...,4;\ \mu=0,1,...,7 \tag{8}$$

将所有的直方图串接为一个序列 $\Re$ 作为最终的人脸描述,本文称其为 HSLGBP.

$$\Re=(H_{0,0,0},...,H_{0,0,m-1},H_{0,1,0},...,H_{0,1,m-1},...,H_{4,7,m-1}) \tag{9}$$

1.4 HSLGBP描述方法的性质分析

HSLGBP 人脸描述方法计算过程简单、直接,无需训练集,且具有对表情、老化、误配准等各种变化鲁棒的特点,分析如下:

1) 对各种图像变化的鲁棒性分析

表情和时间的变化主要表现在面部的局部区域,如吃惊会使眼睛和嘴发生较大变化、皱纹出现在额头和眼角等.基于全局的人脸描述通常对面部的局部变化不鲁棒,但基于局部的人脸描述可以较好地解决这个问题[16].由于本文采用基于区域分析的人脸描述方法,同时直方图序列保留了人脸图像的空间结构信息,因而在保留对人脸图像的整体描述的同时对面部的局部变化也会鲁棒.

2) 对图像误配准的鲁棒性分析

由于特征点定位不精确带来的图像误配准问题往往会引起人脸描述特征的较大“变异”,从而影响最终人脸识别系统的分类识别性能[17].而 HSLGBP 人脸描述方法采用的 Gabor 特征提取方法,其提取的是局部图像特征,即每个 Gabor 特征都是原图像内一定区域范围内的若干个像素共同作用的结果,而不只是单一像素的性质,因此,Gabor 幅值变化更加平缓,即图像中一定范围内的 Gabor 卷积幅值相差不大.这样,当图像中的特征定位出现一定偏差时,对 Gabor 变换的幅值结果影响不大.同时,后端我们进一步采用了描述局部变化的区域直方图,更提高了 HSLGBP 人脸描述方法对图像误配准的鲁棒性.

3) 无须训练,具有良好的推广性

主流的基于统计或者学习策略的人脸识别方法尽管取得了一定的成功,但是,该类方法需要训练数据用于人脸建模,因而存在推广性的问题,即由于测试数据分布与训练数据的分布存在差异,尤其当其分布差异较大时,识别性能下降很大.由于测试数据的未知性,通过扩大训练集的方式并不能很好地解决泛化能力的问题.而基于 HSLGBP 的人脸建模不需要训练集进行训练,且基于 HSLGBP 的匹配也无须训练,因而避免了基于统计学习的人脸建模中的推广能力的问题.

2 基于 HSLGBP 的人脸识别方法

由于HSLGBP是由多个直方图组成的序列,与传统意义上的“人脸描述特征”大相径庭,不能采用欧式距离、余弦夹角之类的相似性度量方法,所以,本文采用直方图的交作为两个直方图序列之间的匹配准则.另一方面,构成 HSLGBP 的每个直方图是从不同的面部区域中提取获得,而不同的面部区域对分类识别的贡献显然应该是不同的,基于此,本文还提出了一种基于 Fisher 判据加权的 HSLGBP 匹配策略,通过对不同的区域直方图赋予不同的权值来提高识别的精度.

2.1 基于直方图交的HSLGBP匹配

直方图交是一种简单而有效的直方图相似度度量方法,其计算方法为

$$\Psi(H^1,H^2)=\sum_{i=1}^{L}\min(h_i^1,h_i^2) \tag{10}$$

其中,H^1 和 H^2 表示两个直方图;L 是直方图 Bin 的数目.这种度量方式是以两直方图 H^1 和 H^2 之间的相同部分的大小来衡量两直方图之间相似性的强弱的.使用这种度量方式,基于 HSLGBP 的人脸描述方法的样本间的相似度可用式(11)计算.

$$S(\Re_1,\Re_2)=\sum_{v=0}^{4}\sum_{\mu=0}^{7}\sum_{r=0}^{m-1}\Psi(H_{v,\mu,r}^1,H_{v,\mu,r}^2) \tag{11}$$

2.2 Fisher准则加权的HSLGBP匹配

如前所述,表情和时间变化对面部的特定区域产生影响.并且,人脸图像中通常作为关键特征的区域,如眼睛、嘴角等本身具有较强的判别能力,而额头、面颊等区域则包含较少的判别信息.因此,对人脸的不同区域赋予不同的权值以反映各个区域的变化规律,将对识别产生积极的影响.据此分析,当进行 HSLGBP 匹配时,我们对不同的空间区域直方图赋予不同的权值.这时,相似度计算公式(11)可重写为式(12)的形式:

$$S'(\Re_1,\Re_2)=\sum_{v=0}^{4}\sum_{\mu=0}^{7}\sum_{r=0}^{m-1}W_{v,\mu,r}\Psi(H_{v,\mu,r}^1,H_{v,\mu,r}^2) \tag{12}$$

这里,$W_{v,\mu,r}$ 是第 g 个 LGBP 图谱的第 r 个区域的权值.本文基于 Fisher 线性判别分析,即类内散度尽可能小而类间散度尽可能大来分析每个区域的权值.但由于人脸识别问题是一个多类分类问题,为应用 Fisher 线性判别准则,本文借鉴文献[18]提出的将多类问题转化为两类问题的“类内差”空间和“类间差”空间的思想,对于 LGBP 图谱中的每个区域,同一类样本的不同样本间的相似度形成“类内相似度”空间,不同类样本的样本间的相似度形成“类间相似度”空间.

假设对于 C 类问题,由上述转化可得:

类内相似度均值为

$$m_{W(v,\mu,r)}=\frac{1}{C}\sum_{i=1}^{C}\frac{2}{N_i(N_i-1)}\sum_{k=2}^{N_i}\sum_{j=1}^{k-1}\Psi(H_{v,\mu,r}^{(i,j)},H_{v,\mu,r}^{(i,k)}) \tag{13}$$

其中,N_i 是第 i 类样本的样本数; $H_{v,\mu,r}^{(i,j)}$ 表示从第 i 类的第 j 个样本的第 (v,μ) 个 LGBP 图谱的第 r 个区域提取的直方图.则类内相似度方差为

$$S_{W(v,\mu,r)}^2=\sum_{i=1}^{C}\sum_{k=2}^{N_i}\sum_{j=1}^{k-1}(\Psi(H_{v,\mu,r}^{(i,j)},H_{v,\mu,r}^{(i,k)})-m_{W(v,\mu,r)})^2 \tag{14}$$

类间相似度均值为

$$m_{B(v,\mu,r)}=\frac{2}{C(C-1)}\sum_{i=1}^{C-1}\sum_{j=i+1}^{C}\frac{1}{N_iN_j}\sum_{k=1}^{N_i}\sum_{l=1}^{N_j}\Psi\left(H_{v,\mu,r}^{(i,k)},H_{v,\mu,r}^{(j,l)}\right) \tag{15}$$

类间相似度方差为

$$S_{B(v,\mu,r)}^2=\sum_{i=1}^{C-1}\sum_{j=i+1}^{C}\sum_{k=1}^{N_i}\sum_{l=1}^{N_j}(\Psi(H_{v,\mu,r}^{(i,k)},H_{v,\mu,r}^{(j,l)})-m_{B(v,\mu,r)})^2 \tag{16}$$

最终,得到 $W_{v,\mu,r}$

$$W_{v,\mu,r}=\frac{(m_{W(v,\mu,r)}-m_{B(v,\mu,r)})^2}{S_{W(v,\mu,r)}^2+S_{B(v,\mu,r)}^2} \tag{17}$$

本文应用上述两种度量方式作为不同直方图序列的相似度度量准则,使用最近邻分类准则进行分类.

3 对比实验与分析

为了评测本文所提方法的性能,我们使用人脸识别研究领域中广泛应用的 FERET 人脸库[2]进行了测试.这是一个规模比较大的人脸数据库,且提供了指定的训练集、原型集(gallery)和测试集(probe).其中,训练集由 429 人的 1 002 幅中性表情和表情变化的图像构成.原型集是由 1 196 人的每人一幅正面图像构成.4 个测试集分别是:fb,fc,Dup.I 和 Dup.II.其中,fb 含有 1 195 幅与原型集同时采集,并且与原型集图像光照相同的表情变化图像;fc 中含有 194 幅与原型集图像采集光照条件不同的人脸图像;在 Dup.I 和 Dup.II 中,分别有 722 幅和 234 幅图像,采集时间距原型集分别为一个月和一年左右.

3.1 评测比对基准算法

为了评价本文提出方法的有效性,我们选择了当前主流的人脸识别方法 Fisherfaces,基于 Gabor 特征的 Fisher 判别作为基准比对算法.同时,与 FERET'97 评测以及文献[11]在 FERET 测试集上的最好结果做比对.

3.1.1 Fisherfaces[5]

Fisherfaces 利用 Fisher 判别准则对数据进行变换,即使得变换后的数据类内散度尽可能小,而类间散度尽可能大.Fisher 线性判别分析是寻找变换矩阵 W,该矩阵是由 $T=S_w^{-1}S_b$ 的按特征值排序的特征向量构成的,其中,S_b 是类间散度矩阵;S_w 是总类内散度矩阵.原始数据 x 在变换矩阵 W 上投影形成新的特征 y,即 $y=W^Tx$.测试集数据与原型集数据分别经过上述变换,然后应用最近邻进行分类.

3.1.2 基于 Gabor 特征的 Fisher 判别[8]

该方法首先对图像进行多尺度、多方向的 Gabor 变换,然后对所得高维数据进行均匀下采样,由于下采样的结果也不能满足类内散度矩阵满秩的条件,所以运用主成分分析对下采样的结果进一步降维,然后进行 Fisher 线性判别分析,本文将其简记为 GFC.由于该方法与本文所提出方法的初始特征相同,选择该方法作比对可以更好地衡量本文所提出方法的性能.同时,也可以从该方法的结果看出基于图像的灰度特征(Fisherfaces)与基于 Gabor 特征的识别结果的差异.

3.1.3 FERET'97[2]评测最好结果

FERET'97[2]评测是对人脸识别具有重要意义的一个评测.首先,它提供了一个公共的人脸数据库,使得不同的人脸识别算法可以此为比对基准;其次,该数据库人脸图像的变化对算法的鲁棒性能提出了很高的要求.通过这个测试,促进了人脸识别向实际应用的发展.其规定的训练集、原型集和测试集也是研究人员测试算法性能常选择的数据库.因此,这个数据库的测试结果也可作为不同算法性能的比对依据.

3.1.4 文献[11]发表的最好测试结果

文献[11]中的结果与文献[2]中的最好结果相比,除 fc 集合以外,均有一定程度的提高.为了评价本文方法与基于图像灰度的 LBP 方法的性能差异,我们选择了文献[11]中的最好结果做比对.

3.2 精确配准实验

3.2.1 区域窗口大小和直方图 bin 对识别率的影响

在本文提出的方法中,区域的大小会对识别性能有一定的影响:如果区域过大,其极端情况即原图像大小,这样无法体现局部区域分析的优势;如果区域过小,其极端情况是像素级分析,这样将对图像的配准敏感性增强.同时,对灰度级进行不同程度的量化也会对识别性能有一定的影响.为观察这些变化对识别性能的影响,我们从 FERET 人脸库的测试集中选出 50 个测试图像,并在 Gallery 中选出其相应的图像,进行识别实验.依据已有的研究成果[19,20],根据 FERET 协议给定的眼睛位置,将图像归一化为 80×88 像素(可参见图 3 给出人脸图像归一化示例).图 2 给出了变化区域窗口大小和对灰度级进行量化的实验结果.

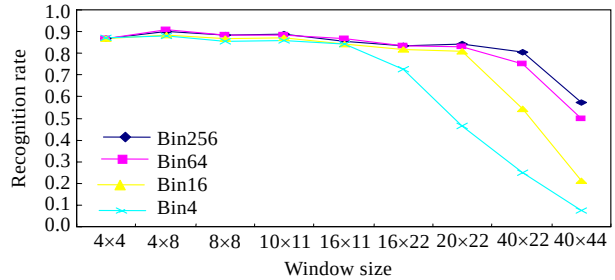


Fig.2 The rank-1 recognition rate of different sizes of region and histogram bin

图 2 不同区域大小和直方图 Bin 时的首选识别率

从图 2 结果可知,当区域窗口相对较小时,由于保留的图像结构信息越多,识别性能越好.但同时,由于区域越小对图像配准和表情等变化越敏感,因此,窗口大小为 4×4 时的识别率低于 4×8;而当窗口变大时,结构信息保留的越少,识别率呈下降趋势.另外,由对灰度级量化的实验结果可知,尽可能地保留灰度级数目可以提高识别性能,这是由于尽可能地逼进原始特征的缘故.但是,这样也会使数据的维数过高.从图 2 可见,当区域窗口较小时,对灰度级进行量化对识别性能影响不是很大.

3.2.2 对表情、光照和时间变化的鲁棒性测试

为测试本文方法的鲁棒性,我们分别在表情、光照和时间变化的 4 个集合上做了测试.为保留更多的人脸图像空间结构信息,同时使特征维数较低,我们选择的区域窗口大小为 4×8 像素,灰度级量化为 16 级.几种方法的实验结果见表 1.

Table 1 The rank-1 recognition rates of different algorithms for the FERET probe sets

表 **1** 不同算法在 FERET 人脸测试集上的首选识别率

Method	fb	fc	Dup.I	Dup.II
Fisherfaces	0.94	0.73	0.55	0.31
GFC	0.95	0.84	0.67	0.61
Best results of FERET'97[2]	0.96	0.82	0.59	0.52
Best results of Ref.[11]	0.97	0.79	0.66	0.64
HSLGBP	0.94	0.97	0.68	0.53
Weighted HSLGBP	0.98	0.97	0.74	0.71

由实验结果可以看出,在 fb 测试集上,几种方法的结果比较接近,而且识别率均在 94%以上.由此可见,当测试图像与原型集图像采集时间相同时,尽管表情发生变化,但识别依然相对容易.从加权 HSLGBP 和 HSLGBP 测试结果可见,对不同的区域赋予不同的权值可以提高识别性能,并且其结果也超过了其他几种算法的结果.

在 fc 上的实验结果相差比较显著:Fisherfaces 的识别率为 73%;GFC 的识别率为 84%.尽管两者都是基于 Fisher 判别分析的,但由于 fc 是相对于原型集发生光照变化的图像集合,而训练集中没有光照变化的图像,因此,Fisherfaces 显现了推广性的问题;而 GFC 的识别率之所以高,是因为 Gabor 特征对光照变化比较鲁棒的原因.基于非统计学习的文献[11]的结果也较差,可见前文分析的基于图像的 LBP 的人脸识别方法具有一定的局限性.而本文 HSLGBP 和加权 HSLGBP 的识别率均为 97%,一方面是由于 Gabor 滤波对光照比较鲁棒,另一方面是因为光照对图像的影响也一定程度地体现在对区域的影响上,例如光照变鼻部影响较大,因此,基于 Gabor 滤波与区域分析的联合作用使得本文的方法在 fc 上取得好的测试结果.

在 Dup.I 和 Dup.II 上,识别结果都有所下降,且 Dup.II 的结果更差.可见测试集图像的采集与原型集图像采集时间间隔越久,人脸图像的变化越大,尤其是局部的变化,这样对识别的影响也就越大.其中在 Dup.II 上,Fisherfaces 仅为 31%;而 GFC 为 61%.这一方面是由于训练集中的样本都是和原型集图像同一时间采集,即没有时间变化的样本;另一方面也说明 Fisherfaces 和 GFC 这种基于全局的分析方法不能很好地表示图像的细节变化.而基于区域的分析,如文献[11]结果为 64%;加权 HSLGBP 的结果为 71%.即可说明基于区域的分析对时间变化具有较好的鲁棒性.此外,在 Dup.II 上,加权 HSLGBP 比 HSLGBP 识别结果高 18%,可以看出,对不同区域赋予不同的权值,可以有效地提高识别性能.

值得指出的是:在 fc 测试中,加权的 HSLGBP 比 HSLGBP 没有提高,我们认为这是由于光照的变化不具有如表情、时间的变化对人脸不同区域影响的规律性(如表情、时间的变化分别特定的区域产生影响).

从测试集为时间变化的图像的测试结果也可以看出,随着时间的变化,尽管人脸的总体表观变化不会很大,但是局部细节的变化对识别的影响还是很大的.

3.3 配准鲁棒测试

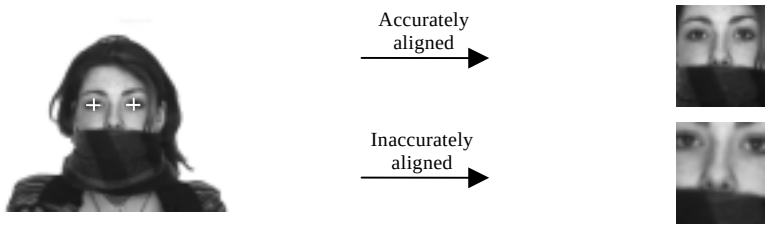
对于实用的人脸识别系统而言,面部特征的精确定位是一个困难的问题.首先,面部特征定位算法在光照、表情甚至遮挡的情况下,会出现一定的偏差;另外,由于图像的成像质量对特征定位也有一定的影响;再者,在用户不配合的非限定条件应用系统中,自动提取的面部特征点往往会有很大的偏差[17].因此,要求人脸识别算法本身对误配准具有较好的鲁棒性.为测试本文提出方法对误配准的鲁棒性,我们采用对手工标定的人眼位置加入随机高斯噪声的策略,获得具有配准误差的人脸测试图像.同时,为观察该方法在其他人脸库上的性能,我们在 AR 人脸库上做了实验.AR 人脸数据库[21]由 126 人的 3 200 多幅正面人脸图像组成.其中每个人 26 幅图像,分两个阶段采集,时间间隔为两周.每阶段采集 13 幅具有不同表情、光照和遮挡的图像.

具体实现过程如下:首先,手工标定人脸图像中双眼位置,如图 3(a)所示;然后,按照双眼位置将图像归一化为具有相同大小且眼睛位置相同的图像.如果按照精确定位归一化人脸图像,将得到如图 3(b)上图所示的图像;如果在归一化时对精确定位加入高斯随机噪声,则得到如图 3(b)下图所示的误配准图像.本文实验采用加入均值为 0、均方差从 0.2~2.4 的高斯随机噪声.为说明本文方法的有效性,我们使用 Fisherfaces 和 GFC 作为评测基准.测试结果如图 4 所示.

由实验结果可见:HSLGBP 随着配准误差的增加,识别精度下降缓慢;而 GFC 和 Fisherfaces 下降都比较快.Fisherfaces 在配准误差的标准差大于 0.8 后识别结果变化较小,是因为本身识别率已经很低.由此可见:本文提出的方法对配准误差鲁棒,当存在配准误差时仍表现出了良好的识别性能.

4 结论及后续工作

本文提出了一种新的基于局部 Gabor 变化直方图序列的人脸描述与识别方法(HSLGBP),该方法首先采用 Gabor 小波滤波器提取人脸图像中各种方向、各种尺度的局部细节变化特征,然后进一步对这些特征进行局部二值编码,并使用局部空间直方图来描述人脸,最终通过直方图匹配完成人脸识别.



(a) Original face image

(a) 原始人脸图像

(b) The top image is accurately aligned and the bottom image is inaccurately aligned

(b) 上图是精确配准图像,下图是有配准误差的图像

Fig.3 Example of face image normalization

图 3 人脸图像归一化示例

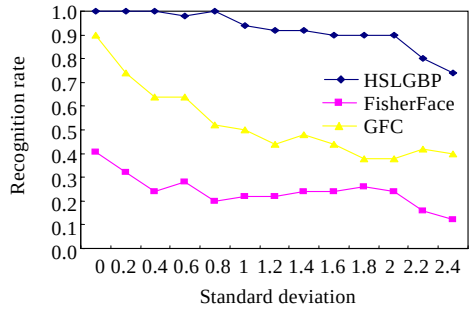


Fig.4 Experiments on robustness to misalignment on AR face database

图 4 在 AR 人脸库上的配准鲁棒测试

本文的方法与传统方法相比,由于采用了 Gabor 小波、局部二值模式、局部空间直方图等不同层次的局部特征提取过程,因而对光照、表情、老化和误配准等问题具有较高的鲁棒性,同时具有较高的识别性能.基于 HSLGBP 的人脸建模方法在本质上不需要在一个训练集上进行训练,并且基于统一权值直方图序列匹配策略时,也无需任何训练集,因而避免了基于统计学习的人脸建模方法必然要面对的推广能力问题.针对能够获得与测试集合相同或相似分布的训练集合的应用环境,我们还提出了在 HSLGBP 匹配阶段采用基于 Fisher 可分性判据的加权方法,以进一步提高识别性能.在 FERET 人脸库的光照、表情和时间变化的测试集上与已知最好结果的对比实验分析表明:本文方法取得了更好的识别性能,完全无需训练过程的统一权值匹配策略取得较好的识别率,采用统计加权策略进一步提高了识别性能.

进一步的工作将集中在如何有效而快速地对直方图序列进行匹配方法的研究上,并探索利用该方法如何适应姿态变化和部分遮挡人脸图像的识别问题.

References:

[1] Phillips PJ, Grother P, Micheals RJ, Blackburn DM, Tabassi E, Bone JM. Face recognition vendor test 2002 results. Evaluation Report, 2003.

[2] Phillips PJ, Syed HM, Rizvi A, Rauss PJ. The FERET evaluation methodology for face-recognition algorithms. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2000,22(10):1090−1104.

[3] Brunelli R, Poggio T. Face recognition: features vs. templates. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1993, 15(10):1042−1053.

[4] Turk M, Pentland A. Face recognition using eigenfaces. In: Negahdaripour S, *et al*., eds. Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. Maui: IEEE Computer Society Press, 1991. 586−591.

[5] Belhumer P, Hespanha P, Kriegman D. Eigenfaecs vs fisherfaces: Recognition using class specific linear projection. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1997,19(7):711−720.

[6] Porat M, Zeevi Y. The generalized Gabor scheme of image representation in biological and machine vision. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1988,10(4):452−468.

[7] Wiskott L, Fellous JM, Kruger N, Malsburg C. Face recognition by elastic bunch graph matching. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1997,19(7):775−779.
[8] Liu CJ, Wechsler H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. IEEE Trans. on Image Processing, 2002,11(4):467−476.
[9] Shan SG. Study on some key issuses in face recognition [Ph.D. Thesis]. Beijing: Institute of Computing Technology, the Chinese Academy of Sciences, 2004 (in Chinese with English abstract).
[10] Vapnik VN, Write; Zhang XG, Trans. The Nature of Statistical Learning Theory. Beijing: Tsinghua University Press, 2000 (in Chinese).
[11] Timo A, Abdenour H, Matti P. Face recognition with local binary patterns. In: Pajdla T, Matas J, eds. Proc. of the European Conf. on Computer Vision. LNCS 3021, Prague: Springer-Verlag, 2004. 469−481.
[12] Lades M, Vorbruggen JC, Buhmann J, Lange J, Malsburg C, Wurtz RP, Konen W. Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. on Computers, 1993,42(3):300−311.
[13] Huang XS, Li SZ, Wang YS. Shape localization based on statistical method using extended local binary pattern. In: Zhang D, Pan ZG, eds. Proc. of the Int'l Conf. on Image and Graphics. Hong Kong: IEEE Computer Society Press, 2004. 184−187.
[14] Jin HL, Liu QS, Lu HQ, Tong XF. Face detection using improved LBP under Bayesian framework. In: Zhang D, Pan ZG, eds. Proc. of the Int'l Conf. on Image and Graphics. Hong Kong: IEEE Computer Society Press, 2004. 306−309.
[15] Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002,24(7):971−987.
[16] Heisele B, Ho P, Wu J, Poggio T. Face recognition: Component-based versus global approaches. Computer Vision and Image Understanding, 2003,91(1−2):6−21.
[17] Shan SG, Chang YZ, Gao W, Cao B. Curse of mis-alignment in face recognition: Problem and a novel mis-alignment learning solution. In: Lee SW, *et al*., eds. Proc. of the IEEE Conf. on Automatic Face and Gesture Recognition. Seoul: IEEE Computer Society Press, 2004. 314−320.
[18] Moghaddam B, Jebara T, Pentland A. Bayesian face recognition. Pattern Recognition, 2000,33(11):1771−1782.
[19] Zhang WC, Shan SG, Zhang HM, Gao W, Chen XL. Multi-Resolution histograms of local variation patterns (MHLVP) for robust face recognition. Kanade T, *et al*., eds. Proc. of the 5th Int'l Conf. on Audio- and Video-Based Biometric Person Authentication. Hilton Rye Town: Springer-Verlag, 2005. 937−944.
[20] Zhang WC, Shan SG, Gao W, Chen XL, Zhang HM. Local Gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition. In: Ma SD, *et al*., eds. Proc. of the IEEE Conf. on Computer Vision. Beijing: IEEE Computer Society Press, 2005. 786−791.
[21] Martinez AM, Benavente R. The AR face database. CVC Technical Report, #24, 1998.

附中文参考文献:

[9] 山世光.人脸识别中若干关键问题的研究[博士学位论文].北京:中国科学院计算技术研究所,2004.
[10] Vapnik VN,著;张学工,译.统计学习理论的本质.北京:清华大学出版社,2000.

张文超(1973－),男,黑龙江双鸭山人,博士生,主要研究领域为图像处理,模式识别,机器学习.

陈杰(1975－),男,博士生,主要研究领域为人脸检测,数字水印,模式识别与图像理解,神经网络.

山世光(1975－),男,博士,副研究员,主要研究为机器视觉,模式识别,机器学习,生物特征识别,人脸识别研究.

陈熙霖(1965－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为多模式人机接口,图像处理,模式识别,计算机视觉,数字电视技术.

张洪明(1976－),男,博士生,主要研究领域为计算机视觉,图像理解,模式识别,机器学习,人机交互界面.

高文(1956－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为计算机视觉,多媒体,图像处理,多模式接口,人工智能,虚拟现实.