软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2015,26(11):2912−2929 [doi: 10.13328/j.cnki.jos.004896]
©中国科学院软件研究所版权所有.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于差值局部方向模式的人脸特征表示*

李照奎[1], 丁立新[2], 王 岩[1], 何进荣[3], 丁国辉[1]

[1](沈阳航空航天大学 计算机学院,辽宁 沈阳 110136)

[2](软件工程国家重点实验室(武汉大学),湖北 武汉 430072)

[3](西北农林科技大学 信息工程学院,陕西 杨凌 712100)

通讯作者: 李照奎, E-mail: lzk@sau.edu.cn, http://www.sau.edu.cn

摘 要: 提出一种基于差值局部方向模式的人脸特征表示方法(difference local directional pattern,简称 DLDP):首先,通过 Kirsch 掩模卷积运算,为每个像素计算 8 个方向的边缘响应值;然后,计算 8 个相邻边缘响应值的强度差,前 k 个最突出的强度差对应的方向编码为 1,其他方向编码为 0,形成一个 8 位二进制数表示对应的 DLDP 模式;此外,针对高分辨率的 Kirsch 掩模单纯考虑方向性而没有考虑像素位置权重的问题,提出相应的掩模权值设计方法;最后,把每幅图像划分成多个不重叠的局部图像块,通过统计图像块上不同 DLDP 模式个数生成相应的子直方图,所有子直方图被串联起来表示一幅人脸图像.实验结果表明,该方法在光照、表情、姿态和遮挡方面获得了较好的结果,尤其针对遮挡情况,表现更为突出.

关键词: 差值局部方向模式;Kirsch 掩模;特征表示;人脸识别

中图法分类号: TP391

中文引用格式: 李照奎,丁立新,王岩,何进荣,丁国辉.基于差值局部方向模式的人脸特征表示.软件学报,2015,26(11):2912−2929. http://www.jos.org.cn/1000-9825/4896.htm

英文引用格式: Li ZK, Ding LX, Wang Y, He JR, Ding GH. Face feature representation based on difference local directional pattern. Ruan Jian Xue Bao/Journal of Software, 2015,26(11):2912−2929 (in Chinese). http://www.jos.org.cn/1000-9825/4896.htm

Face Feature Representation Based on Difference Local Directional Pattern

LI Zhao-Kui[1], DING Li-Xin[2], WANG Yan[1], HE Jin-Rong[3], DING Guo-Hui[1]

[1](School of Computer, Shenyang Aerospace University, Shenyang 110136, China)

[2](State Key Laboratory of Software Engineering (Wuhan University), Wuhan 430072, China)

[3](College of Information Engineering, Northwest A & F University, Yangling 712100, China)

Abstract: A face feature representation method based on difference local directional pattern (DLDP) is proposed. Firstly, each pixel of every facial image sub-block gains eight edge response values by convolving the local neighborhood with eight Kirsch masks. Then, the difference of each pair of neighboring edge response values is calculated to form eight new difference directions. The top k difference response values are selected and the corresponding directional bits are set to 1. The remaining (8−k) bits are set to 0, thus forming the binary expression of a difference local direction pattern. In addition, high-resolution Kirsch masks only consider directions but ignore the weight values of each pixel location. DLDP proposes a design method of weight values. Finally, the sub-histogram is calculated by accumulating the number of different DLDP of image blocks. All sub-histograms of an image are concatenated into a new face descriptor. Experimental results show that DLDP achieves state-of-the-art performance for difficult problems such as expression, illumination and occlusion-robust face recognition in most cases. Especially, DLDP gets better results for occlusion problem.

Key words: difference local directional pattern (DLDP); Kirsch mask; feature representation; face recognition

* 基金项目: 国家自然科学基金(61170185,61303016); 辽宁省教育厅科学研究一般项目(L2015403); 珠海市重点实验室科技攻关项目(2012D0501990026); 沈阳航空航天大学校博士启动金(15YB05)

收稿时间: 2015-05-26; 修改时间: 2015-07-14; 定稿时间: 2015-08-26

© 中国科学院软件研究所 http://www.jos.org.cn

在过去几十年间,生物度量在诸如监控、人机接口、司法和安全标识等许多领域扮演着非常重要的角色[1].生物度量识别是一种自动进行个体识别的方法,它通过比较生物体的生理或行为特性来进行识别,这些生理或行为特性往往被转换成一个特征向量.生理特性主要包括人脸、指纹、掌纹、手指形状、手掌形状、虹膜、视网膜以及声音等[2].人脸图像可以通过廉价的照相设备在一定的距离之外方便地获取,它不需要用户的主动参与,也没有任何健康风险[3].因此,人脸识别系统往往更容易被用户接受,并且得到了研究人员的持续关注.

常用的方法有主成分分析(principal component analysis,简称 PCA)[4]、独立成分分析(independent component analysis,简称 ICA)[5]、线性判别分析 linear discriminant analysis,简称 LDA)[6]、二维主成分分析(two-dimensional principal component analysis,简称 2DPCA)[7]、等度规特征映射(isometric feature mapping,简称 ISOMAP)[8]、局部线性嵌入(local linear embedding,简称 LLE)[9]、局部保持投影(locality preserving projections,简称 LPP)[10]、近邻保持嵌入(neighborhood preserving embedding,简称 NPE)[11]、边界费舍尔分析(marginal fisher analysis,简称 MFA)[12]、拉普拉斯核维数约简方法(Laplacian kernel dimensionality reduction method,简称 LK-RDRM)[13]、梯度方向与欧拉映射(gradient orientations and Euler mapping,简称 GOGM)[14]、拉普拉斯特征值映射(Laplacian eigenmaps,简称 LE)[15]等.基于局部的表示方法主要从图像的局部块进行考虑,提取的是局部特征,常见的方法有 Gabor 特征[16]、局部二进制模式((local binary pattern,简称 LBP)[17]、局部三进制模式(local ternary pattern,简称 LTP)[18]、局部 Gabor 二进制模式(local Gabor binary pattern,简称 LGBP[19]、中心对称局部二进制模式(center-symmetric local binary pattern ,简称 CS-LBP)[20]、Gabor 体局部二进制模式(Gabor volume local binary pattern,简称 GV-LBP[21]、局部方向模式(local directional pattern,简称 LDP)[22]和增强局部方向模式(enhanced local directional patterns,简称 ELDP)[23]等.一般地,全局表示方法对光照、姿态、表情以及遮挡等复杂变化比较敏感.越来越多的研究表明:局部表示方法由于刻画了局部更细节的纹理信息,对光照、姿态、表情以及遮挡等复杂变化往往表现得更为鲁棒[19,21−23].为了提高全局特征对光照、姿态、表情及遮挡等复杂变化的鲁棒性,文献[24]提出了一类基于图像梯度方向的子空间学习方法(image gradient orientations,简称 IGO).IGO 方法首先把图像转换为梯度图像,然后在梯度空间[25]上进行全局维数约简.梯度图像更多地考虑了局部像素之间的关系,本质上是在获取图像的局部结构信息.与直接在图像强度值上处理相比,梯度图像的计算可以获得更好的鲁棒性,但是在遮挡处理方面,IGO 方法仍然未能达到理想的效果.

LBP[17]作为一种简单而有效的局部人脸表示方法,已经引起众多研究人员的注意.LBP 主要被用在人脸识别和纹理分析上,它依据中心像素和其邻域像素之差的正负来进行二进制编码,具有旋转不变性.它对单调光照变化比较鲁棒,但对非单调光照变化和随机噪声比较敏感.Jabid 等人[22]提出一种基于局部方向模式(LDP)的更鲁棒的人脸特征表示方法.针对某个中心像素,LDP 表示方法并不直接比较中心像素和相邻像素的强度值,而是通过 8 个 Krisch 掩模,在某个邻域内为中心像素计算 8 个方向的边缘响应值,然后把绝对值最大的 k 个边缘响应值所对应的二进制位设置为 1,剩下的 $8-k$ 个二进制位设置为 0.Zhong 等人[23]提出一种基于 ELDP 的人脸表示方法,ELDP 只是考虑最强和次强的边缘响应值,它对边缘响应值的选取主要考虑边缘响应值的方向性,它把值最大和次大的边缘响应值(不考虑绝对值)所对应的位置,通过 8 进制方式进行编码.与稀疏采样的 LBP 相比,LDP 和 ELDP 实质是对局部邻域的稠密采样.从统计观点看,稀疏采样会丢失更多的邻域信息,降低方法的正确性,而 Kirsch 掩模的采用会获得更鲁棒的效果.因此当有噪声存在时,LDP 和 ELDP 会比 LBP 表现得更稳定.ELDP 只对最大和次大的边缘响应值进行编码,编码结构更加简单.此外,与 LDP 相比,ELDP 考虑了边缘响应值的方向性.尽管 LDP 和 ELDP 采用稠密采样,获取了更多的邻域信息,但它们对遮挡情况仍然比较敏感.

本文提出一种基于差值局部方向模式(difference local directional pattern,简称 DLDP)的人脸表示方法.LDP 和 ELDP 通过 8 个 Kirsch 掩模和图像的卷积结果进行编码,一个重要的问题是,LDP 和 ELDP 并没有考虑近邻边缘响应值之间的强度变化.通常,强度变化越大,往往表示两个方向之间存在的边缘特征或对角线特征越突出.DLDP 通过计算近邻边缘响应值之间的强度变化,本质上是在获取方向之间更丰富的细节特征.强度变化越大,往往表示两个方向之间的细节特征越突出,更突出的细节特征往往具有更强的判别力.另一个重要的问题是:LDP 和 ELDP 受到原始 Kirsch 掩模的制约,只考虑 3×3 掩模,传统的高分辨率 Kirsch 掩模只是单纯考虑方向

© 中国科学院软件研究所 http://www.jos.org.cn

性(如在正东方向上,所有元素都设置为 1),并没有考虑不同像素位置的权重.通常认为,距离参照点越近的权重越大,距离参照点越远的权重越小.通过分析 3×3 的 Kirsch 掩模所蕴含的约束机制,提出高分辨率 Kirsch 掩模的权值设计方法.此外,LDP 和 ELDP 主要考虑在强度空间上进行编码,DLDP 将在梯度空间进行编码.显然,梯度空间比强度空间包含更多的信息.比如,梯度空间可以保留像素之间的内在关系,而强度空间则不能.梯度空间保留这些内在关系,实质上揭示了图像的潜在结构信息,而这些结构信息往往具有更强的判别性.为了保证梯度计算更稳定,DLDP 首先使用高斯方法平滑图像.高斯平滑可以减弱噪声的影响,而梯度空间的计算会使 DLDP 对光照变化更鲁棒.实验结果显示:DLDP 在光照、表情、姿态和遮挡方面获取了较好的结果,尤其针对遮挡情况,DLDP 表现更为突出.

本文第 1 节介绍相关研究.第 2 节提出相应的 DLDP 方法.第 3 节进行 Kirsch 掩模分析与加权设计.第 4 节描述相应的人脸表示与识别方法.实验结果与分析展示在第 5 节.最后,在第 6 节进行总结.

1 相关研究

1.1 LBP

LBP[17]主要依据中心像素与其邻居像素之差的正负来进行二进制编码.针对图像中的每个像素,在以该像素为中心的一个局部邻域内,通过比较相邻像素和中心像素的大小来对中心像素进行二进制编码.当邻域像素与中心像素的差值为正值或为 0 时,相应的二进制位被编码为 1;而当邻居像素与中心像素的差值为负值时,相应的二进制位被编码为 0.

图 1 展示了一个 3×3 的图像块中,像素的坐标位置及其对应的像素强度值.中心像素(x_c,y_c)一共有 $P(=8)$个局部邻居像素,中心像素的强度值为 g_c,所有局部邻居像素的强度值分别为$\{g_0,g_1,\ldots,g_7\}$.LBP 的具体编码方法如下:

$$LBP_{P,R}(x_c,y_c)=\sum_{i=0}^{P-1}s(g_i-g_c)2^i,s(x)=\begin{cases}1, & x\geqslant 0\\ 0, & x<0\end{cases} \tag{1}$$

这里,g_c 表示中心像素的强度值,g_p 表示以中心像素为圆心、半径为 R 的圆上的 P 个邻居像素的强度值.

(x_c-1,y_c-1)	(x_c,y_c-1)	(x_c+1,y_c-1)
(x_c-1,y_c)	(x_c,y_c)	(x_c+1,y_c)
(x_c-1,y_c+1)	(x_c,y_c+1)	(x_c+1,y_c+1)

(a) 像素坐标

g_3	g_2	g_1
g_4	g_c	g_0
g_5	g_6	g_7

(b) 像素强度值

Fig.1 Pixel coordinates and intensity values in a 3×3 image patch

图 1 3×3 图像块中的像素坐标及强度值

利用公式(1)可以为每个像素获得一个 8 位的二进制编码(在 3×3 邻域内),因此,最多可以获得 128(2^8)种计算结果,也就是常说的 128 种模式.但 Ojala 等人[26]观察到:在大量图像的编码中,只有有限的模式会频繁地出现,这些模式被称作 uniform LBP.可以把二进制编码看作一个循环数据位序列,这些统一模式具有一个共同的特点,即,在循环序列中,相邻数据位从 0 到 1 或从 1 到 0 的变化次数最少.一般把变化次数小于等于 2 的模式称作 uniform LBP,这样的模式一共有 58 种.为了充分利用所有出现的模式,LBP 编码中往往把剩余的 70(128−58)种模式当作一种单独的模式,因此在 LBP 算法中,用到的模式总共有 59 种.尽管 LBP 在纹理分析和人脸识别上已经取得了极大的成功,但 LBP 仍然对对非单调光照变化和随机噪声比较敏感.

1.2 LDP

局部方向模式(LDP)[22]也为图像的每个像素分配一个 8 位二进制编码,它主要通过比较某个像素在不同方向上的边缘响应值(edge response value,简称 erv)来获取模式.LDP 通过 Kirsch 掩模为中心像素计算 8 个不同的

© 中国科学院软件研究所 http://www.jos.org.cn

方向响应值.图 2 展示了 8 个不同方向的 Kirsch 掩模 M_i,i=0,1,…,7.8 个 Kirsch 掩模分别表示 8 个方向所对应的 8 条边,东向、西向、南向和北向分别对应直线边,东北向、西北向、西南向和东南向分别对应折线边.应用 8 个掩模对图像块进行卷积运算,我们可以获得中心像素的 8 个边缘响应值 $ERV=\{erv_0,erv_1,\ldots,erv_7\}$,计算方法见公式(2):

$$erv_i=I*M_i \tag{2}$$

这里,*表示卷积运算,I 表示输入图像,M_i 表示第 i 个掩模,erv_i 表示第 i 个边缘响应值.

−3	−3	5
−3	0	5
−3	−3	5

East (M_0)

−3	5	5
−3	0	5
−3	−3	−3

North East (M_1)

5	5	5
−3	0	−3
−3	−3	−3

North (M_2)

5	5	−3
5	0	−3
−3	−3	−3

North West (M_3)

5	−3	−3
5	0	−3
5	−3	−3

West (M_4)

−3	−3	−3
5	0	−3
5	5	−3

South West (M_5)

−3	−3	−3
−3	0	−3
5	5	5

South (M_6)

−3	−3	−3
−3	0	5
−3	5	5

South East (M_7)

Fig.2 Kirsch edge response masks in eight directions (3×3)

图 2 在 8 个方向上的 Kirsch 边缘响应掩模(3×3)

每个边缘响应值分别体现了在不同方向上边的重要性,在所有方向上的边缘响应值的重要性并不等同.LDP 模式主要寻找前 k 个最大的反应值的绝对值,并把相应方向的二进制编码值设置为 1,而其他的(8−k)个方向设置为 0.具体编码方法见公式(3):

$$LDP(x_c,y_c)=\sum_{i=0}^{7}s(|erv_i|-\Psi)2^i,s(x)=\begin{cases}1, & x\geqslant 0\\ 0, & x<0\end{cases} \tag{3}$$

这里,Ψ表示第 k 个最大的边缘响应值的绝对值$|erv_k|$,$\Psi=\mathrm{Max}^{k\mathrm{th}}(ERV)$;$ERV=\{|erv_0|,|erv_1|,\ldots,|erv_k|,\ldots,|erv_7|\}$.

利用 Kirsch 掩模对图像进行卷积运算,我们将为每个像素分配一个 8 位的二进制编码.LDP 总共有 C_8^k 种不同的模式,如果 k 取 3,则共有 56 种模式.

1.3 ELDP

LDP 使用边缘响应值的绝对值进行编码,Zhong 等人[23]认为,原始的边缘响应值的正负意味着某个方向上两个相反的趋势(上升或下降),可能包含更多的信息,因此,他们提出一种增强局部方向模式(enhanced local directional pattern,简称 ELDP).在 ELDP 中,首先通过 LDP 中的方法计算 8 个边缘响应值 $ERV=\{erv_0,erv_1,\ldots,erv_7\}$;然后,通过边缘响应值 erv_i 的下标值 i 来表示 erv_i 的方向值,因此,8 个方向正好可以分别用 8 个八进制编码来表示(0,1,2,3,4,5,6,7 分别对应 000,001,011,100,101,110,111).ELDP 利用最大边缘响应值和次大边缘响应值的方向值进行编码,因此,ELDP 编码可以用一个双八进制数$(ab)_8$来表示,这个双八进制数可以通过公式(4)转换成十进制数:

$$ELDP(x_c,y_c)=a\times 8^1+b\times 8^0 \tag{4}$$

在这里,a 是最大边缘响应值 m_a 的索引下标,b 是次大边缘响应值 m_b 的索引下标,其中,$a,b\in[0,7]$.ELDP 总共有 $P_8^2=8\times 7=56$ 种不同的模式,由于 8 个位置取不同的八进制数,因此是排列而不是组合.与 LDP 编码相比,ELDP 编码只是对前两个最大边缘响应值对应的位置进行编码,显然,编码结果要相对简洁很多.此外,ELDP 考虑边缘响应值的方向性,因此大部分情况选择的可能都是正响应值对应的方向.而通常,更小的负值也可能意味着更突出的图像边缘细节.

© 中国科学院软件研究所 http://www.jos.org.cn

2 DLDP

2.1 DLDP编码

LDP 通过 8 个 Kirsch 掩模和图像的卷积结果进行编码,在 LDP 编码中,分别为每个参照像素分配一个 8 位二进制数来表示 8 个方向上的边缘响应强弱.ELDP 通过考虑边缘响应值的方向性,提出相应的简化版编码方法.特别需要注意的是:LDP 和 ELDP 都没有考虑近邻边缘响应值之间的强度变化,强度变化越大,往往表示两个方向之间存在更突出的边缘特征.另一个问题是,LDP 和 ELDP 受到原始 Kirsch 掩模的制约,只是考虑 3×3 掩模.我们将在第 3 节把 Kirsch 掩模扩展到更高的分辨率,并且依据原始 3×3 掩模的权值设计约束,调整高分辨率中的掩模权值.

此外,LDP 和 ELDP 主要考虑在强度空间上进行编码,DLDP 将在梯度空间进行编码.显然,梯度空间比强度空间包含更多的信息.比如,梯度空间可以保留像素之间的内在关系,而强度空间则不能.梯度空间保留这些内在关系实质上揭示了图像的潜在结构信息,而这些结构信息往往具有更强的判别性.为了保证梯度计算更稳定,DLDP 首先使用高斯方法平滑图像.高斯平滑可以减弱噪声的影响,梯度空间的计算会使 DLDP 对光照变化更鲁棒.

针对一幅变换后的梯度图像,我们通过计算近邻边缘响应值之间的差值 $derv$ 来表示近邻边缘响应值之间的强度变化,然后计算相应的编码.具体编码方法见公式(5):

$$DLDP(x_c,y_c)=\sum_{i=0}^{7}s(|derv_i|-\Xi)2^i,\ s(x)=\begin{cases}1, & x\geqslant 0\\ 0, & x<0\end{cases} \tag{5}$$

这里,

- $derv_i$ 表示近邻边缘响应值的差值,我们称作差值边缘响应值:

  $derv_0=erv_0-erv_7, derv_1=erv_1-erv_0,\ldots,derv_7=erv_7-erv_6$;

- Ξ表示第 k 个最大的差值边缘响应值的绝对值$|derv_k|$:

  $\Xi=\mathrm{Max}^{kth}(DERV); DERV=\{|derv_0|,|derv_1|,\ldots,|derv_7|\}$.

由于采用了近邻边缘响应值之间的差值,因此,本算法被称为差值局部方向模式,图 3 展示了具体的 DLDP 编码实例.DLDP 和 LDP 类似,也有 C_8^k 种不同的模式,如果 k 取 3,则共有 56 种模式.

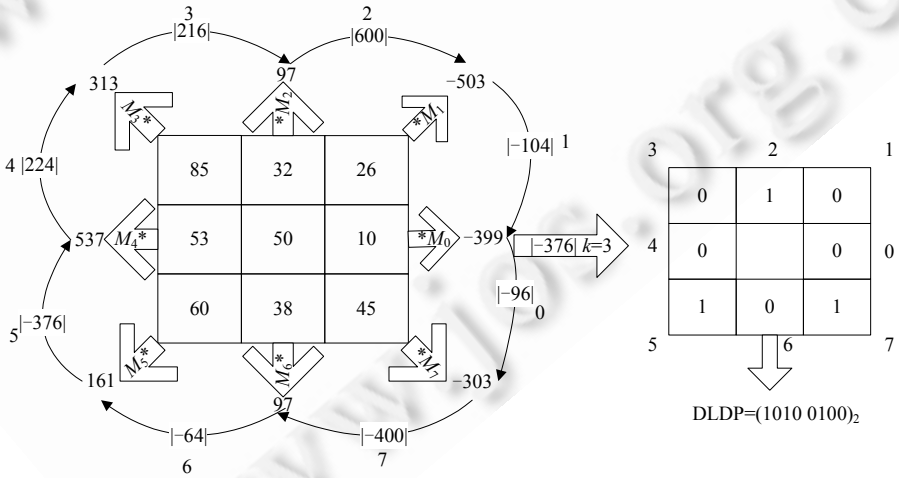


Fig.3 An example of calculating DLDP code when k=3

图 3 当 k 取 3 时,DLDP 的编码实例

图 4 展示了 AR 数据中某个人的部分图像和对应的 DLDP 图像.图 4(a)对应非遮挡的原始图像,图 4(b)是其相应的 DLDP 编码图像.图 4(c)对应遮挡的原始图像,图 4(d)是其相应的 DLDP 编码图像.从图 4 中可以看出,

© 中国科学院软件研究所 http://www.jos.org.cn

DLDP 编码图像提取了更多边缘细节.通常意义上,边缘细节往往具有更强的判别性.

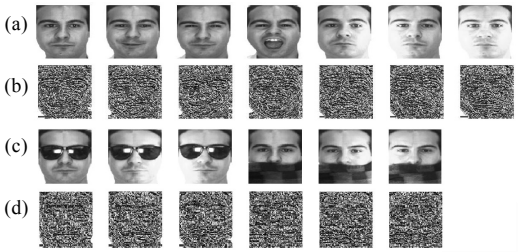


Fig.4 Original images and DLDP code images

图 4 原始图像和 DLDP 编码图像

2.2 DLDP的稳定性和可靠性

LBP 和 LDP 的主要区别在于编码时对邻居集的采样方式,针对某个中心像素在某一方向上的编码,LBP 往往只选择邻居集的 1 个像素或 3 个像素,属于稀疏采样;而 LDP 在计算某一方向的编码时,采用的是稠密采样,即,选择邻居内的所有像素.从统计观点来看,Kirsch 掩模的采用会获得更鲁棒的效果;而稀疏采样会丢失更多的邻域信息,降低方法的正确性.因此,当有轻微噪声存在时,LDP,ELDP 和 DLDP 会比 LBP 表现得更稳定.

LDP 可以在局部邻域内获取更多的边缘特征:如果在某个方向上的边缘响应值为正值,则说明在该方向上存在由黑到白的边缘特征;如果在该方向上边缘响应值为负,则说明在该方向上存在有白到黑的边缘特征.正值或负值的绝对值越大,边缘特征的反差越大,也就是说,边缘特征越明显.ELDP 主要考虑 LDP 计算的边缘响应值的符号,但问题是 ELDP 考虑的只是边缘响应值最大的两个,因此很有可能错过负值最小的边缘响应值,也就是说,会错过一些比较明显的由白到黑的边缘特征,这可能会导致 ELDP 在某些情况下要弱于 LDP.

LDP 和 ELDP 并没有考虑近邻边缘响应值之间的强度变化,DLDP 通过计算近邻边缘响应值之间的强度变化来提取特征.通常认为:相邻的边缘响应值之差的绝对值越大,则说明在两个方向之间存在的特征越明显.这极有可能是两个方向的边缘响应值一个为正一个为负,显然一个从黑到白,一个从白到黑,此时,应该得到的是所谓的对角线特征.当然,也有可能是两个方向具有相同的符号,但二者相差很大,这说明在两个方向之间存在比较明显的边缘特征.因此,与 LDP 和 ELDP 相比,DLDP 在局部邻域中获取了更丰富的细节特征,强度变化越大,往往表示两个方向之间的细节特征更突出,更突出的细节特征往往具有更强的判别力.

下面通过图 5 中的例子来验证,图 5 展示了 3 个局部 3×3 图像块,图 5(a)是原始图像块,图 5(b)是相应的高斯白噪声图像块,图 5(c)是相应的遮挡图像块.

85	32	26
53	50	10
60	38	45

LBP=$(00111000)_2$
LDP=$(00010011)_2$
ELDP=$(43)_8$
DLDP=$(10100100)_2$

(a) 原始图像和编码

90	35	22
49	55	11
64	40	41

LBP=$(00101000)_2$
LDP=$(00010011)_2$
ELDP=$(43)_8$
DLDP=$(10100100)_2$

(b) 相应的高斯白噪声图像和编码

81	80	80
53	50	10
60	38	45

LBP=$(00111110)_2$
LDP=$(10001100)_2$
ELDP=$(23)_8$
DLDP=$(10100100)_2$

(c) 相应的遮挡图像和编码

Fig.5 Stability of LBP, LDP, ELDP and ELDP

图 5 LBP,LDP,ELDP 和 DLDP 的稳定性

从图 5(a)和图 5(b)可以看出,由于增加了白噪声,LBP 的第 5 位从 1 变为 0,因此,LBP 模式从统一模式变成

© 中国科学院软件研究所 http://www.jos.org.cn

非统一模式.而由于 LDP,ELDP 和 DLDP 通过稠密采样,因此,即使在噪声情况下,两者也都产生了同样的编码模式.在图 5(c)中,我们把第 1 行设置为近似相等的像素以表示遮挡情况,此时,LBP,LDP 和 ELDP 的模式都有别于图 5(a)和图 5(b)中对应的 LBP,LDP 和 ELDP 模式.但由于 DLDP 考虑了局部近邻方向上边缘响应值的差异性,因而即使在遮挡情况下,DLDP 也产生了同样的编码模式.这充分说明了 DLDP 编码对遮挡的鲁棒性.

3 Kirsch 掩模分析与加权设计

3.1 Kirsch掩模分析

图 6 展示了第 i(=0,1,…,7)个 Kirsch 掩模, M_i^8 为中心位置权值, $M_i^j\,(j=0,1,...,7)$ 为 M_i^8 的 8 个邻居位置权值,结合图 2 可以看出,Kirsch 掩模主要满足如下 3 个约束:

(1) 对于任意一个 Kirsch 掩模 M_i,中心位置 M_i^8 的权值为 0;

(2) 对于任意一个 Kirsch 掩模 M_i,中心位置的所有邻居位置的权值之和为 0,即有 $\sum_{j=0}^{7} M_i^j = 0$;

(3) 任意一个掩模中有 3 个较大权值 w^l,5 个较小权值 w^s,较大权值所在的方向代表掩模的方向.

8 个 Kirsch 掩模主要通过权值的设置来刻画 8 个不同的方向,如,为了获取东北方向掩模(如图 7 所示),则为东北方向上 3 个连续邻居位置(实心圆形)设置较大的权值 w^l,而为其他 5 个连续邻居位置(空心圆形)设置较小的权值 w^s.由于要满足约束(2),因此对于东北方向掩模中,中心位置的所有邻居位置的权值应满足第 3 个约束,则有如下公式:

$$3\times w^l + 5\times w^s = 0 \Rightarrow \frac{w^l}{w^s} = \frac{5}{-3} \tag{6}$$

因此我们可以得出 w^l 和 w^s 的比值关系,原始的 Kirsch 掩模实际上选取了最小整数权值 5 和−3,当然,也可以取 5/3 和−1,或者 10 和−6.

M_i^3	M_i^2	M_i^1
M_i^4	M_i^8	M_i^0
M_i^5	M_i^6	M_i^7

Fig.6 The ith Kirsch mask

图 6 第 i 个 Kirsch 掩模

○	●	●
○	0	●
○	○	○

Fig.7 Northeast directional mask

图 7 东北方向掩模

3.2 加权设计

本节主要针对不同分辨率下的掩模进行加权设计.图 8 显示了原始的 5×5 Kirsch 掩模.这里的 Kirsch 掩模只是单纯考虑方向性(如在正东方向上,所有元素都设置为 1),并没有考虑不同像素位置的权重.通常认为:距离参照点(即中心点 M_i^8)越近的权重越大,距离参照点越远的权重越小.

对于任一分辨率为 $m\times m$(m 为大于 1 的奇数)的掩模,总共有 $\frac{m-1}{2}$ 层,τ_r 为第 r 层邻居个数,$M^{(r),j}$ 表示掩模的第 r 层的第 j 个邻居,第 r 层的邻居从 0 排列到 τ_r-1,为了满足约束(2),则有:

$$\sum_{j=0}^{\tau_1-1} M^{(1),j} + \sum_{j=0}^{\tau_2-1} M^{(2),j} + ... + \sum_{j=0}^{\tau_r-1} M^{(r),j} + ... + \sum_{j=0}^{\tau_{(m-1/2)}-1} M^{((m-1)/2),j} = 0 \tag{7}$$

对于公式(7),我们只需保证每层的权值之和为 0 即可.由公式(6),可以类似地计算出通用的掩模中心位置的每层邻居的权值比值.公式(8)是第 r 层的较大权值和较小权值的比值:

$$(2r+1)\times w_r^l + (\tau_r - (2r+1))\times w_r^s = 0 \Rightarrow \frac{w_r^l}{w_r^s} = \frac{\tau_r - (2r+1)}{-(2r+1)} \tag{8}$$

© 中国科学院软件研究所 http://www.jos.org.cn

其中,τ_r 为第 r 层邻居个数, w_r^l, w_r^s 分别对应第 r 层中设置的较大权值和较小权值.

−1	−1	0	1	1
−1	−1	0	1	1
−1	−1	0	1	1
−1	−1	0	1	1
−1	−1	0	1	1

East (M_0)

0	1	1	1	1
−1	0	1	1	1
−1	−1	0	1	1
−1	−1	−1	0	1
−1	−1	−1	−1	0

North East (M_1)

1	1	1	1	1
1	1	1	1	1
0	0	0	0	0
−1	−1	−1	−1	1
−1	−1	−1	−1	−1

North (M_2)

1	1	1	1	0
1	1	1	0	−1
1	1	0	−1	−1
1	0	−1	−1	−1
0	−1	−1	−1	−1

North West (M_3)

1	1	0	−1	−1
1	1	0	−1	−1
1	1	0	−1	−1
1	1	0	−1	−1
1	1	0	−1	−1

West (M_4)

0	−1	−1	−1	−1
1	0	−1	−1	−1
1	1	0	−1	−1
1	1	1	0	−1
1	1	1	1	0

South West (M_5)

−1	−1	−1	−1	−1
−1	−1	−1	−1	−1
0	0	0	0	0
1	1	1	1	1
1	1	1	1	1

South (M_6)

−1	−1	−1	−1	0
−1	−1	−1	0	1
−1	−1	0	1	1
−1	0	1	1	1
0	1	1	1	1

South East (M_7)

Fig.8 Kirsch edge response masks in eight directions (5×5)

图 8 在 8 个方向上的 Kirsch 边缘响应掩模(5×5)

下面以 9×9(m=9)掩模为例,计算各层权值.9×9 掩模总共有 4 层,从第 1 层~第 4 层的邻居个数分别为 8,16,24 和 32.根据公式(8),分别设置如下权值:

- 第 1 层:处于该层某方向上的邻居位置的较大权值为 5,其他邻居位置的较小权值为−3;
- 第 2 层:处于该层某方向上的邻居位置的较大权值为 11,其他邻居位置的较小权值为−5;
- 第 3 层:处于该层某方向上的邻居位置的较大权值为 17,其他邻居位置的较小权值为−7;
- 第 4 层:处于该层某方向上的邻居位置的较大权值为 23,其他邻居位置的较小权值为−9.

图 9 展示了不同层的邻居和相应的权值(对应东北方向),图 9(a)~图 9(d)分别对应 1~4 层.

−3	5	5
−3	0	5
−3	−3	−3

(a)

−5	−5	11	11	11
−5				11
−5		0		11
−5				−5
−5	−5	−5	−5	−5

(b)

−7	−7	−7	17	17	17	17
−7						17
−7						17
−7			0			17
−7						−7
−7						−7
−7	−7	−7	−7	−7	−7	−7

(c)

−9	−9	−9	−9	23	23	23	23	23
−9								23
−9								23
−9								23
−9				0				23
−9								−9
−9								−9
−9								−9
−9	−9	−9	−9	−9	−9	−9	−9	−9

(d)

Fig.9 Neighbors and the corresponding weight values under different layers (northeast direction)

图 9 不同层的邻居和相应的权值(对应东北方向)

根据图 9 中的权值设置,可以得出完整的 9×9 Kirsch 掩模(对应东北方向),具体如图 10(d)所示,图 10(a)~图 10(c)分别是 3×3,5×5 和 7×7 掩模.但通常认为:距离参照点(即中心点)越近的权重越大(取绝对值),距离参照点越远的权重越小(取绝对值),图 10 的权值设置显然不尽合理.当然,相同距离的位置,权值的设置遵照处于方向上的权值大于剩余的位置的权值(取绝对值).公式(9)用来计算对第 r 层中的权值进行放大的倍数,σ为倍数调整参数.σ越小,内层与外层权值变化就越小;相反,σ越大,内、外层权值变化就越大.

$$\lambda_r = \left|\frac{w_{r+1}^l}{w_r^s}\right| + \sigma, \sigma > 0 \tag{9}$$

© 中国科学院软件研究所 http://www.jos.org.cn

−3	−3	5
−3	0	5
−3	−3	5

(a)

−5	−5	11	11	11
−5	−3	5	5	11
−5	−3	0	5	11
−5	−3	−3	−3	−5
−5	−5	−5	−5	−5

(b)

−7	−7	−7	17	17	17	17
−7	−5	−5	11	11	11	17
−7	−5	−3	5	5	11	17
−7	−5	−3	0	5	11	17
−7	−5	−3	−3	−3	−5	−7
−7	−5	−5	−5	−5	−5	−7
−7	−7	−7	−7	−7	−7	−7

(c)

−9	−9	−9	−9	23	23	23	23	23
−9	−7	−7	−7	17	17	17	17	23
−9	−7	−5	−5	11	11	11	17	23
−9	−7	−5	−3	5	5	11	17	23
−9	−7	−5	−3	0	5	11	17	23
−9	−7	−5	−3	−3	−3	−5	−7	−9
−9	−7	−5	−5	−5	−5	−5	−7	−9
−9	−7	−7	−7	−7	−7	−7	−7	−9
−9	−9	−9	−9	−9	−9	−9	−9	−9

(d)

Fig.10 Krisch masks under different resolutions (northeast direction)

图 10 不同分辨率下的 Kirsch 掩模(对应东北方向)

对于 9×9 掩模,相对于中心位置总共有 4 层,其中,前 3 层需要调整权值,各层权值调整参数计算如下:

- 第 1 层:该层权值矩阵调整倍数为 $\lambda_1=\left|\frac{11}{-3}\right|+\sigma,\sigma>0$;
- 第 2 层:该层权值矩阵调整倍数为 $\lambda_2=\left|\frac{17}{-5}\right|+\sigma,\sigma>0$;
- 第 3 层:该层权值矩阵调整倍数为 $\lambda_3=\left|\frac{23}{-7}\right|+\sigma,\sigma>0$.

设 $M^{(r)}$表示第 r 层的所有邻居权值,为了保证内层权值大于外层权值,则需要乘上所有大于等于 r 层的的调整倍数.具体计算见公式(10):

$$M^{(r)}=\lambda_{\max-1}\times\lambda_{\max-2}\times\ldots\times\lambda_r\times M^{(r)},\max=\frac{m-1}{2},r=1,\ldots,\max-1 \tag{10}$$

max 是 $m\times m$ 掩模中的最大层数,最外层不需要调整权值.

图 11 展示了调整权值后的 5×5 Kirsch 掩模(对应东北方向),此时调整参数 σ 取 1/3,调整倍数 $\lambda_1=\left|\frac{11}{-3}\right|+\frac{1}{3}=4$.经过调整后的 5×5 Kirsch 不但满足所有 3 个约束,而且满足距离参照点越远权值越小的思想.

−5	−5	11	11	11
−5	−12	20	20	11
−5	−12	0	20	11
−5	−12	−12	−12	−5
−5	−5	−5	−5	−5

Fig.11 Weighted 5×5 Kirsch mask (northeast direction)

图 11 调整权值后的 5×5 Kirsch 掩模(对应东北方向)

4 人脸表示与识别

4.1 人脸表示

给定一张图像 I,通过 DLDP 操作后,我们可以获得一张编码图像 I_{DL},在编码图像时,把边缘响应值最大的 k=3 个方向编码为 1,把剩余的 5 个方向编码为 0,最终可以获得 56(C_8^3)种不同的编码模式 dp_i,i=1,2,…,56.为了获取图像中丰富的局部纹理信息(例如边、点和角点等),通常采用直方图来统计局部图像块中的编码模式.接下来详述图像的直方图表示方法.

针对一张经过 DLDP 编码后的图像 I_{DL},我们把它划分成 L 个非重叠的局部图像块 $I_{DL}=[b^1,\ldots,b^j,\ldots,b^L]$,

© 中国科学院软件研究所 http://www.jos.org.cn

$b^j \in \Re^{w\times w}$,其中,w=4,8.对于编码图像中的每个局部块 b^j,可以通过如下公式计算其子直方图:

$$H^j(dp_i)=\sum_{(x,y)} P(b^j(x,y)=dp_i) \text{ where } P(A)=\begin{cases}1, & A=\text{true}\\ 0, & A=\text{false}\end{cases} \tag{11}$$

在子直方图的计算中,$H^j(dp_i)$通过累加第 j 个子块中属于 dp_i 模式的像素的个数来计算.子直方图 $H^j(dp)$实际对应了一个长度为 56 的列向量 h^j,向量中的每个元素表示 j 个子块属于每个每个模式的像素个数.对于一幅编码图像 I_{DL},串联所有的子直方图,以形成一个更具判别性的直方图 H_{DL}:

$$H_{DL}=\prod_{j=1}^{L} H^j \tag{12}$$

这里,$\prod$表示串联操作,L 表示一张图像划分的块的个数,串联后的直方图对应一个长度为 $56\times L$ 的列向量 h_{DL}. H_{DL} 在空间上组合了整张图像的统计模式,实际上扮演了全局特征脸的角色.因此,对于一张图像 I,最终用列向量 h_{DL} 来表示.图 12 展示了用 DLDP 进行人脸表示的流程.

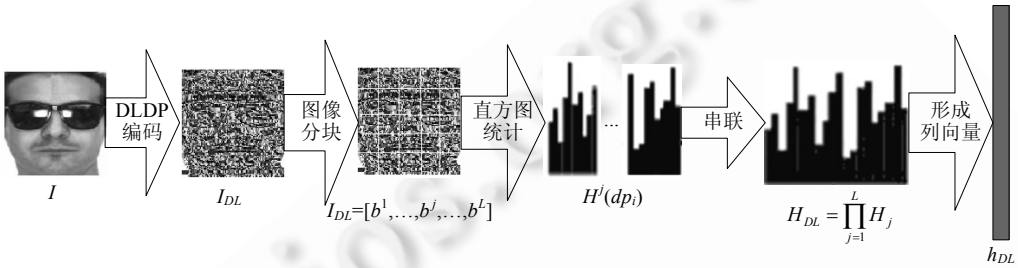


Fig.12 Face representation using DLDP

图 12 用 DLDP 进行人脸表示

4.2 人脸识别

在人脸识别阶段,首先抽取所有候选脸和测试脸的直方图特征向量 h_{DL},然后通过计算测试脸和候选脸的特征向量之间的相似性来判别测试脸属于哪一张候选脸.直方图向量之间的相似性测度通常采用 Chi-square 统计方法.给定两张脸的直方图特征向量 $h^1_{DL}\in\Re^{56L}$ 和 $h^2_{DL}\in\Re^{56L}$,$h^1_{DL}(i)$ 表示 h^1_{DL} 的第 i(=1,2,…,56L)维,则 h^1_{DL} 和 h^2_{DL} 之间的 Chi-square 统计测度定义为

$$\chi^2(h^1_{DL},h^2_{DL})=\sum_{i=1}^{56L}\frac{(h^1_{DL}(i)-h^2_{DL}(i))^2}{h^1_{DL}(i)+h^2_{DL}(i)} \tag{13}$$

利用公式(13)计算测试脸与所有候选脸之间的 Chi-square 测度值,选择具有最小 Chi-square 统计测度值的候选脸与测试脸匹配.

5 实 验

本节主要评估我们提出的 DLDP 算法,并且与一些时新的算法进行比较.我们将在 4 个公开的人脸数据库(AR[27],Extended Yale B[28],CMU PIE[29,30]以及 FERET[31])上执行相关的实验,我们的实验涉及更广泛的人脸变化,主要有人脸表情变化(AR,CMU PIE 和 FERET)、光照变化(AR,Extended Yale B,CMU PIE 和 FERET)、原有遮挡(AR)、人为遮挡(AR,Extended Yale B,CMU PIE 和 FERET)、年龄(FERET)、轻度的姿态变化(CMU PIE 和 FERET).在实验中,AR 图像的分辨率设置为 100×90,Extended Yale B 图像的分辨率设置为 96×84,CMU PIE 图像和 FERET 图像的分辨率都被设置为 64×64.

实验中所有涉及 LDA 的方法(如 LDA,IGO-LDA),为了避免类内散度矩阵 S_w 的奇异性,都采用带正则化参数的优化方法$(S_w+\varepsilon I)^{-1}S_bQ=Q\Lambda$.这里,$\varepsilon$是一个小的正则化常量(取 10^{-3} 以下),I 是一个单位矩阵.

对于 LBP 特征,采样点为 8 个,采样半径为 2,在 AR 上的图像块设置为 16×15,在其他数据库上的图像块设

© 中国科学院软件研究所 http://www.jos.org.cn

置为 8×8,每块对应一个 59 维的直方图.对于 LTP 特征,在 AR 上的图像块设置为 16×15,在其他数据库上图像块设置为 8×8,每个图像块对应一个 128 维的直方图.

对于 LDP,编码为 1 的个数 k=3;在 AR,Extended Yale B 和 CMU PIE 上的图像块设置为 4×4,在 FERET 上的图像块设置为 8×8,每个图像块都对应一个 56 维的直方图.

对于 ELDP 特征,把最大值和次大值对应的位置编码为 8 进制数,在 AR,Extended Yale B 和 CMU PIE 上的图像块设置为 4×4,在 FERET 上的图像块设置为 8×8,每个图像块都对应一个 56 维的直方图.

对于 DLDP,在 AR 中编码为 1 的个数 k=2(此时,每个图像块都对应一个 28 维的直方图);在 Extended Yale B, CMU PIE 和 FERET 中编码为 1 的个数 k=3(此时,每个图像块都对应一个 56 维的直方图);所有数据库中掩模分辨率设置为 5×5,加权倍数调整参数σ设置为 1.在 AR,Extended Yale B 和 CMU PIE 上图像块设置为 4×4,在 FERET 上图像块设置为 8×8.

为了保证公平性,采用简单的 1 近邻分类器(1-NN)进行分类,距离度量方式采用余弦测度.而针对 LBP,LTP, LDP,ELDP,DLDP,距离度量方式都采用 Chi-square 统计测度.

为了评价算法的鲁棒性,我们增加了一系列人为的遮挡实验.在这些遮挡实验中,主要用不同分辨率的恐龙头部图像,在测试图像的任意位置上进行遮挡.针对 AR 和 Extended Yale B,分别使用分辨率为 55×55 的恐龙头部图像;针对 CMU PIE 和 FERET,使用分辨率为 45×45 的恐龙头部图像.

5.1 AR数据库

AR 人脸库[27]包括 4 000 多张正面人脸图像,包括 126 个人(70 个男性,56 个女性).每个人有 26 张图像,被分为两部分:第 1 部分每个人共有 13 张图像,从 1~13 进行编号,其中包括不同的人脸表情(1~4)、光照变化(5~7)以及在不同光照条件下的不同遮挡(8~13);第 2 部分与第 1 部分类似,也有 13 张图像,分别对应表情、光照和遮挡.图 13 展示了用在实验中的图像样例,第 1 行的 7 张图像来自于第 1 部分,第 2 行的 7 张图像来自于第 2 部分的非遮挡图像,第 3 行的 7 张图像是对第 2 行图像通过人为遮挡后的图像(恐龙头部遮挡),第 4 行的 6 张图像是来自于第 2 部分的系统遮挡图像(墨镜遮挡和围巾遮挡).

Fig.13 Face images of the same subject taken from the AR database

图 13 来自 AR 的同一个人的多张图像

针对 AR 数据库,我们共做了 9 个实验,表 1 提供了每个实验详细的训练集和测试集信息.从表 1 可以看出,所有实验的训练集被分为 3 类:第 1 部分的多个非遮挡图像(编号 1~7),对应实验 1~实验 3;第 1 部分的较少非遮挡图像(编号 1,2),对应实验 4~实验 6;第 1 部分的单个非遮挡图像(编号 1),对应实验 7~实验 9.所有实验的测试集也被分为 3 类:第 2 部分的非遮挡图像(编号 1~7);对第 2 部分的非遮挡图像(编号 1~7)进行人为遮挡(恐龙头部遮挡);第 2 部分的系统遮挡图像(编号 8~13,墨镜遮挡和围巾遮挡).表 2 展示了实验 1~实验 9 的相关实验结果.

从表 2 可以看出,DLDP 算法在所有情况下都获得了最好的识别率.实验 1 主要针对非遮挡情况,PCA 的识别率较差,LDA 的识别率要高于 PCA,但并不令人满意.IGO-PCA,IGO-LDA,LBP,LTP,LDP,ELDP 和 DLEP 都取得了较好的识别效果,然而 DLDP 的识别效果是最好的.实验 2 和实验 3 是遮挡情况下的识别效果,显然,PCA,LDA

© 中国科学院软件研究所 http://www.jos.org.cn

的识别效果急速下降,说明全局特征方法对遮挡情况非常敏感.IGO-PCA,IGO-LDA 和 LBP,LTP 的识别效果要好于 PCA 和 LDA.值得注意的是,LDP,ELDP 和 DLDP 都获得了较好的识别效果,这充分说明了稠密邻居采样的作用.但 DLDP 获得更好的识别效果,与排名第二的算法相比,分别高出 2.42 个百分点和 2.33 个百分点.

Table 1 All experiment configurations on the AR database

表 1 AR 人脸库中所有实验的详细信息

实验	训练集(来自第1部分)	测试集(来自第2部分)
1	1~7 (非遮挡)	1~7 (非遮挡)
2	1~7 (非遮挡)	1~7 (用恐龙头部图像在任意位置进行人为遮挡)
3	1~7 (非遮挡)	8~13 (系统遮挡:墨镜遮挡和围巾遮挡)
4	1,2 (非遮挡)	1~7 (非遮挡)
5	1,2 (非遮挡)	1~7 (用恐龙头部图像在任意位置进行人为遮挡)
6	1,2 (非遮挡)	8~13 (系统遮挡:墨镜遮挡和围巾遮挡)
7	1 (非遮挡)	1~7 (非遮挡)
8	1 (非遮挡)	1~7 (用恐龙头部图像在任意位置进行人为遮挡)
9	1 (非遮挡)	8~13 (系统遮挡:墨镜遮挡和围巾遮挡)

Table 2 Recognition rates (%) of different methods on the AR database (Exp.1~Exp.9)

表 2 在 AR 人脸库中,不同识别方法的识别率(%)(实验 1~实验 9)

方法	实验 1	实验 2	实验 3	实验 4	实验 5	实验 6	实验 7	实验 8	实验 9
PCA	72.43	25.86	22	48.71	18.86	14.5	48.14	19.42	14.83
LDA	86.86	27.43	36.83	53.85	19.29	22.17	49.43	17.43	15.17
IGO-PCA	97.29	83.29	82.5	86.86	73.57	71.83	84.42	72.71	70.17
IGO-LDA	98.14	85.71	83.5	87.29	75.57	74.83	85.21	71.13	71.31
LBP	94.28	84.72	80.67	87	63.43	58.33	85.71	58.29	59.33
LTP	93.71	84.86	80.67	85.86	62.14	58.67	84.14	57.57	60
LDP	98.29	95	94.5	89.58	80.57	86.5	85.86	77.43	84
ELDP	97.29	95.29	91	88.43	82.86	81.67	86.29	80.29	80.33
DLDP	**98.86**	**97.71**	**96.83**	**93.57**	**87**	**91.67**	**93**	**82.86**	**90.17**

实验 4 不涉及遮挡问题,PCA,LDA 表现不够理想.这表明当训练样本个数变少时,全局特征算法表现得较差.IGO-PCA,IGO-LDA 获得了不错的识别效果,说明梯度空间的采用可以获取更多的结构信息,而结构信息显然更有助于识别率的提升.LBP,LTP,LDP 和 ELDP 也获得了不错的识别效果,说明局部特征对光照和表情比较鲁棒.由于 DLDP 在梯度空间上提取局部的细节特征,因此取得了最好的识别结果,与识别率第二的 LDP 相比,提高了 3.99 个百分点.在实验 5 和实验 6 中,以正常图像作为训练,以遮挡图像作为测试,是更具挑战性的实验.在实验 5 中,我们通过恐龙头部图像在测试图像任意位置进行遮挡;在实验 6 中,我们选用数据库本身提供的遮挡图像(墨镜遮挡和面纱遮挡),遮挡位置比较固定.

从实验 5 和实验 6 的结果可以看出:PCA,LDA 表现得更差,IGO-PCA,IGO-LDA,LBP 和 LTP 的识别效果要好于 PCA 和 LDA,但仍不能令人满意.LDP,ELDP 和 DLDP 都取得了不错的识别效果,其中,我们提出的 DLDP 表现更好,与排名第二的算法相比(实验 5 中的 ELDP 和实验 6 中的 LDP),分别高出 4.14 个百分点和 5.17 个百分点.DLDP 获得较好的识别效果,主要原因有 3 点:第一,DLDP 通过计算近邻边缘响应值之间的强度变化,本质上是在获取方向之间的更丰富的细节特征,强度变化越大,往往表示两个方向之间的细节特征越突出,更突出的细节特征往往具有更强的判别力;第二,DLDP 通过分析 Kirsch 3×3 掩模所蕴含的约束机制,提出了高分辨率 Kirsch 掩模的权值设计方法;第三,由于 DLDP 在梯度空间上计算编码,因此对光照更鲁棒,此外,在梯度计算时采用高斯函数进行图像平滑,因此减弱了噪声的影响.

在实验 7~实验 9 中,每个人的训练样本个数减少为 1 张图像,此时,传统的全局特征表示方法 PCA 和 LDA 的识别效果明显下降,这充分说明全局特征过于依赖于训练样本的个数.在遮挡情况下(实验 8 和实验 9),PCA 和 LDA 识别效果更差.在非遮挡情况下(实验 7),IGO-PCA,IGO-LDA,LBP,LTP,LDP,ELDP 和 DLDP 也获得了不错的效果,这与这些特征蕴含了局部信息有关.其中,我们提出的 DLDP 方法获得最好的识别效果,与排名第二的 ELDP 相比,提高了 6.71 个百分点.与非遮挡情况相比,在遮挡情况下(实验 8 和实验 9),IGO-PCA,IGO-LDA,LBP,

© 中国科学院软件研究所 http://www.jos.org.cn

LTP,LDP,ELDP 和 DLDP 的识别效果都有所下降,但其中 IGO-PCA,IGO-LDA,LBP 和 LTP 下降得更为明显,LDP, ELDP 和 DLDP 下降得相对较少.与排名第二的算法相比(实验 8 中的 ELDP 和实验 9 中的 LDP),DLDP 分别高出 2.57 个和 6.17 个百分点.

5.2 Extended Yale B数据库

Extended Yale B 人脸库[28]包含 2 414 张正面人脸图像,包括 38 个人,每个人有大约 64 张图像,这些图像是在不同光照条件下采集的.尽管图像的姿态变化不大,但极度的光照条件对大多数人脸识别方法来说仍然是一个不小的挑战.所有标记为 P00 的前脸图像被应用在我们的实验中,图 14 展示了用在实验中的一个人的多张人脸图像,其中,第 1 行是原始的非遮挡图像,第 2 行是相应的人为遮挡图像(用恐龙头部图像在第 1 行图像的任意位置进行遮挡).

Fig.14 Partial face images of the same subject taken from the Extended Yale B database

图 14 来自 Extended Yale B 的同一个人的多张图像

我们一共执行两组实验:

- 在实验 1 中,我们为每个人随机选择 N_{train} 个图像作为训练集,剩下的图像作为测试集.这里,N_{train} 分别取 4,8, 12,16,20 和 24.对于 N_{train} 的每种选择,分别随机生成 10 种不同的训练集进行实验;
- 实验 2 和实验 1 类似,唯一的不同是在实验 2 中,所有的测试图像通过恐龙头部图像块进行人为遮挡. 两组实验中所有方法的平均识别率被展示在表 3 中.

Table 3 Average recognition rates (%) of different methods on the Extended Yale B database (Exp.1~Exp.2)

表 3 在 Extended Yale B 人脸库中,不同识别方法的平均识别率(%) (实验 1、实验 2)

方法	实验 1(非遮挡)						实验 2(恐龙头部遮挡)					
	4	8	12	16	20	24	4	8	12	16	20	24
PCA	33.69	46.02	58.94	63.07	71.28	73.44	23.43	31.85	39.89	43.13	46.61	47.87
LDA	74.40	86.21	92.70	96.17	97.88	98.6	50.57	60.76	71.09	73.98	79.87	79.69
IGO-PCA	97.66	99.35	99.42	99.52	99.58	99.61	88.68	91.13	88.87	86.21	80.41	78.09
IGO-LDA	**98.98**	**99.41**	**99.53**	**99.61**	99.62	99.65	92.16	96.1	97.54	98.70	99.23	99.57
LBP	62.60	77.58	88.15	93.58	95.34	96.07	61	76.21	87.29	92.03	93.66	94.61
LTP	62.29	77.54	87.95	93.69	95.28	96.07	61.37	76.75	87.39	91.47	93.53	94.87
LDP	86.60	93.70	97.19	98.01	98.91	99.2	83.24	92.89	96.32	97.31	98.32	98.4
ELDP	77.06	87.96	94.63	97.62	98.49	99	76.3	88.15	94.38	97.11	98.51	98.6
DLDP	94.91	98.75	99.29	99.61	**99.70**	**99.66**	**93.06**	**97.63**	**98.52**	**99.31**	**99.45**	**99.67**

从表 3 可以看出:

- 在非遮挡情况下(实验 1),PCA 的识别效果比较差,LBP 和 LTP 获得了优于以上算法的识别结果,LDA 获得了相对较好的识别效果,基于稠密邻域采样的 LDP,ELDP,DLDP 都获得了更好的识别效果.尽管当每人的训练样本个数较少时(选取 4,8 和 12),DLDP 的识别率要低于 IGO-PCA,IGO-LDA,但随着训练样本的增加,DLDP 逐渐取得了最好的识别率,并且识别率几乎都达到 100%.
- 在遮挡情况下(实验 2),PCA,LDA 等传统全局特征方法表现得更差,LBP,LTP,LDP,ELDP,DLDP,IGO-PCA 和 IGO-LDA 都有不同程度的下降,但 DLDP 明显下降得更少,在所有情况都获得了最优的识别率,这充分说明 DLDP 算法对遮挡的鲁棒性.

5.3 CMU PIE数据库

CMU PIE[29,30]人脸数据库包含 41 368 张人脸图像,包括 68 个人,每个人的图像有 13 种不同的姿态、43 种

© 中国科学院软件研究所 http://www.jos.org.cn

不同的光照条件和 4 种不同的表情.本实验选择 PIE 数据库的子集[30],该子集中每人包含 50 张人脸图像,这 50 张图像分属 5 种不同的姿态,每种姿态包含 10 张图像.图 15 展示了用在我们实验中的一个样本的多张图像,其中,第 1 行是原始的非遮挡图像,第 2 行是相应的人为遮挡图像(用恐龙头部图像在第 1 行图像的任意位置进行遮挡).

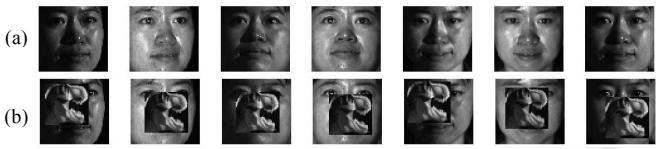


Fig.15 Partial face images of the same subject taken from the CMU PIE database

图 15 来自 CMP PIE 的同一个人的多张图像

我们执行两组实验:

- 在实验 1 中,我们为每个人随机选择 N_{train} 个图像作为训练集,剩下的图像作为测试集.这里,N_{train} 分别取 5,10,15,20 和 25.对于 N_{train} 的每种选择,分别随机生成 10 种不同的训练集进行实验;
- 实验 2 和实验 1 类似,唯一的不同是在实验 2 中,所有的测试图像通过恐龙头部图像块进行人为遮挡.

两组实验中,所有方法的平均识别率被展示在表 4 中.由于 PIE 数据库涉及到姿态、光照和表情,显然更具挑战性.

Table 4 Average recognition rates (%) of different methods on the CMU-PIE database (Exp.1~Exp.2)

表 4 在 CMU PIE 人脸库中,不同识别方法的平均识别率(%) (实验 1、实验 2)

方法	实验 1(非遮挡)					实验 2(恐龙头部遮挡)				
	5	10	15	20	25	5	10	15	20	25
PCA	31.63	47.24	58.03	66.72	73.41	14.93	19.30	25.04	27.06	29.88
LDA	68.24	83.79	87.52	90.59	90.82	16.73	22.5	26.97	27.06	35.76
IGO-PCA	72.91	87.35	90.42	93.28	92	54.64	70.4	75.38	77.89	78
IGO-LDA	**80.95**	**93.38**	**95.38**	96.16	96.76	52.94	65.63	72.73	72.94	75.47
LBP	62.48	80.74	87.71	92.7	95.12	51.86	66.84	77.94	83.98	82.76
LTP	63.63	80.4	87.69	92.75	96.71	51.60	66.58	77.86	83.28	82.35
LDP	66.87	82.24	87.86	93.97	92.65	54.31	70.37	79.79	85.44	90.29
ELDP	61.70	79.6	87.02	91.96	93	54.15	71.21	80.55	87.79	91.29
DLDP	76.76	88.9	93.23	**96.58**	**97.18**	**58.33**	**74.15**	**82.31**	**88.48**	**91.71**

从表 4 可以看出:

- 在非遮挡情况下(实验 1),PCA 的识别效果明显较差,LDA,LBP,LTP,LDP 和 ELDP 的识别效果优于以上算法,IGO-PCA,IGO-LDA 和 DLDP 获得了相对较好的识别效果.在非遮挡情况下,当训练样本较少时,DLDP 并没有获得最优识别率,但随着训练样本的增加,DLDP 的识别率不仅赶上了排名第一的 IGO-LDA,而且在训练样本为 20 和 25 时超过了它.
- 在遮挡情况下(实验 2),无论训练样本多少,DLDP 都获得了最优识别率.而在实验 1 中表现突出的 IGO-LDA 算法,在遮挡情况下识别率明显下降.这充分表明 DLDP 算法对遮挡情况的鲁棒性.

5.4 FERET数据库

FERET 数据库[31]由美国国防部通过 DARPA(defense advanced research projects agency)计划发起,它已经成为一个测试和评估人脸识别算法的标准数据库.我们提出的算法在 FERET 的一个子集上进行测试,这个子集包括 1 400 张图像、200 个人(每个人有 7 张图像).该子集涉及人脸表情、光照和姿态的变化.每个人的 7 张图像包括 3 张正脸图像和 4 张侧面图像,3 张正脸图像包括光照、表情和时间(分别被标记为‘ba’,‘bj’,‘bk’),4 张侧面图像具有具有±15°和±25°的姿态(分别被标记为‘bd’,‘be’,‘bf’,‘bg’)[30].图 16 展示了用在我们实验中的一个样本的多张图像,其中,第 1 行是原始的非遮挡图像,第 2 行是相应的人为遮挡图像(用恐龙头部图像在第 1 行图像的

© 中国科学院软件研究所 http://www.jos.org.cn

任意位置进行遮挡).

Fig.16 Face images of the same subject taken from the FERET database

图 16 来自 FERET 的同一个人的多张图像

我们执行两组实验:

- 在实验 1 中,我们从 FERET 数据库中为每个人选择 2 张图像作为训练集,剩余的作为测试.6 种不同的训练集组合情况分别是‘ba’和‘bd’、‘ba’和‘be’、‘ba’和‘bf’、‘ba’和‘bg’、‘ba’和‘bj’、‘ba’和‘bk’.我们分别计算了每种算法在这 6 种情况下的识别率,也计算了这 6 种情况下的平均识别率.
- 实验 2 和实验 1 类似,唯一的不同是在实验 2 中,所有的测试图像通过恐龙头部图像块进行人为遮挡.

两组实验的测试结果展示在表 5 中.

Table 5 Recognition rates (%) of different methods on the FERET database (Exp.1~Exp.2)

表 5 在 FERET 人脸库中,不同识别方法的识别率(%)(实验 1、实验 2)

方法	实验 1(非遮挡)							实验 2(恐龙头部遮挡)						
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Mean	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Mean
PCA	55.7	52.3	52.1	54.5	40	37	48.6	12.4	12.4	10.7	10.2	12.8	6.5	10.83
LDA	34.7	37.1	29.4	29.9	22	20	28.85	6.1	7.7	5.5	5.6	5.4	5.6	5.98
IGO-PCA	54.5	55	40.1	46.4	29.10	28.9	42.33	28.7	30.8	23.3	26	16.5	16.2	23.58
IGO-LDA	52.9	53.7	40	47.1	34	33.7	43.57	24.3	24.6	22.8	24.1	19	20.8	22.6
LBP	82	82	78.9	78.9	70.7	71.8	77.38	47.4	46.6	40.9	42.8	34	37.6	41.55
LTP	81.7	81.7	79	78.8	70.9	71.8	77.32	45.6	44.6	39.7	41.1	33.6	35.5	40.02
LDP	77.9	77.7	73.7	73.8	63.5	66	72.1	45.5	46.2	40.9	41.1	30.4	35.2	39.88
ELDP	77.4	77	74.6	74	64.2	66.2	72.23	52.6	52.8	49.4	49.3	40	43.1	47.87
DLDP	**83.3**	**83.2**	**81.2**	**81**	**72.7**	**74**	**79.23**	**61.9**	**61.2**	**57.5**	**57.6**	**45.9**	**49.3**	**55.57**

从表 5 可以看出:

- 在非遮挡情况下(实验 1),DLDP 在所有情况都获得了最佳识别率,DLDP 的平均识别率比排名第二的 LBP 高出 1.85 个百分点,效果并不是很明显;
- 在遮挡情况下(实验 2),DLDP 同样获取了最佳识别率.值得注意的是:在实验 2 中,DLDP 的平均识别率比排名第二的 ELDP 高出 7.7 个百分点,比 LBP 高出 14.02 个百分点.这充分说明了 DLDP 对遮挡的鲁棒性.

5.5 参数分析

本节主要讨论 DLDP 的相关参数设置.在 DLDP 中,主要涉及 3 个参数:选择编码为 1 的个数 k、掩模分辨率 $m\times m$ 和加权倍数调整参数σ.接下来,在以上提到的 4 个数据库上,评估 DLDP 方法的参数设置对识别率的影响.我们首先分析参数 k 的取值和对应的模式情况,表 6 展示了 k 在取不同的值时,所拥有的模式数量.参数 k 不可能取 8,因为此时只有一种模式.当 k 取 1 或 7 时,只有 8 种模式存在,显然模式个数过少,识别效果不可能好.因此,在这里我们只讨论 k 取 2,3,4,5,6 时的情况.

Table 6 Number of patterns under different k

表 6 不同的 k 所对应的模式数量

k	1	2	3	4	5	6	7	8
模式数量	8	28	56	70	56	28	8	1

在 AR 数据库上,以每个人的第 1 部分的 1 张中性表情图像作为训练,以第 2 部分的 7 张图像作为测试.针

© 中国科学院软件研究所 http://www.jos.org.cn

对 Extended Yale B 数据库,我们为每个人随机选择 8 个图像作为训练集,剩下的图像作为测试集,一共进行 10 种不同的随机划分.针对 CMU PIE 数据库,我们为每个人随机选择 10 个图像作为训练集,剩下的图像作为测试集,同样进行 10 种不同的随机划分.然后,对得到的 10 个识别率取平均值,通过比较最终的平均识别率来获取最佳参数配置.针对 FERET 数据库,我们针对第 6 种情况进行测试.由于 3×3 掩模不涉及参数σ,因此,在任意的σ下都取同样的值.表 7 展示了在 4 个人脸数据库 AR,Extended Yale B,CMU PIE 和 FERET 上的实验结果.

Table 7 Recognition rates (%) of DLDP with different parameters on four face databases

表 7 在 4 个人脸库中,不同参数下 DLDP 的识别率(%)

		σ															
k	$m\times m$	AR				Extended Yale B				CMU PIE				FERET			
		0.5	1	1.5	2	0.5	1	1.5	2	0.5	1	1.5	2	0.5	1	1.5	2
2	3×3	89.86	89.86	89.86	89.86	97.58	97.58	97.58	97.58	84.53	84.53	84.53	84.53	66.2	66.2	66.2	66.2
	5×5	92.43	**93**	92.86	92	98.25	98.23	98.25	98.25	87.82	87.86	87.85	87.84	72.1	71.9	71.4	71.3
	7×7	91.57	92	91.42	91.43	97.86	97.72	97.87	97.97	86.1	86.31	86.23	86.12	71.4	71.5	71.3	71.2
3	3×3	89.57	89.57	89.57	89.57	98.20	98.20	98.20	98.20	85.21	85.21	85.21	85.21	68.8	68.8	68.8	68.8
	5×5	91.43	91	91.57	91.14	98.72	**98.75**	98.72	98.72	88.57	**88.9**	88.81	88.76	73.4	**74**	73.7	73.6
	7×7	90	89.71	90	90.23	98.63	98.43	98.44	98.44	86.23	86.54	86.49	86.15	73.4	73.4	72.9	72.7
4	3×3	87.86	87.57	87.57	87.57	98.20	98.20	98.20	98.20	83.17	83.17	83.17	83.17	65.8	65.8	65.8	65.8
	5×5	89.14	88.57	89.14	89	98.72	98.72	98.72	98.72	87.67	87.84	87.76	87.15	72.3	72.5	70.8	71.2
	7×7	87	88.57	88.43	88.43	98.48	98.58	98.58	98.53	85	86.29	86.62	86.31	71.3	71.7	70.8	70.8
5	3×3	87.14	87.14	87.14	87.14	98	98	98	98	82.72	82.72	82.72	82.72	64.7	64.7	64.7	64.7
	5×5	89.29	89	88.86	88.71	98.1	98.1	98.1	98.1	86.52	86.89	86.79	86.63	72.8	72.7	72.3	71.8
	7x7	89.43	88.86	88.57	88.14	98.1	98.16	98.19	98.25	84.37	84.43	84.21	83.79	71.5	71.2	71.4	71
6	3×3	87.86	87.86	87.86	87.86	98.43	98.43	98.43	98.43	81.42	81.42	81.42	81.42	61.8	61.8	61.8	61.8
	5×5	89	88.86	88.29	88.43	98	98	98.68	98	85.45	85.78	85.54	84.43	68.5	69.5	68.5	68.8
	7×7	87.86	87.14	87.57	87.43	97.78	97.63	97.68	98.78	83.33	83.56	83.43	82.71	67.9	68.7	68.4	68.2

从表 7 我们有如下观察:

(1) 无论哪一个数据库的实验结果,当掩模分辨率取 3×3 时,其识别效果要弱于分辨率取 5×5 和 7×7,这充分说明了加权调整的作用.其中,掩模分辨率取 5×5 时获得更好的识别率.

(2) 与 k 取 4,5,6 和 7 相比,当 k 取 2 或 3 时,DLDP 获得更高的识别率,其中,当 k 取 2 时,AR 数据库获得了更好的识别率;但当 k 取 3 时,DLDP 获得第二好的识别率.当 k 取 3 时,Extended Yale B,CMU PIE 和 FERET 数据库获得了更好的识别率.

(3) 当掩模分辨率相同、k 取值相同时,不同的σ下的识别率相差不大;但当σ取 1 时,在大部分情况下都能获得更好的识别率.

综上分析,针对不同人脸数据库,通常选取的最佳参数配置为:k 取 3,掩模分辨率取 5×5,加权倍数调整参数σ取 1.

6 结束语

本文提出一种基于差值局部方向模式的人脸特征表示方法 DLDP.与 LDP 和 ELDP 等考虑在强度空间上进行编码不同,DLDP 主要在梯度空间进行编码.显然,梯度空间比强度空间包含更多的结构信息.然后,DLDP 通过计算近邻边缘响应值之间的差值来获取近邻方向之间更丰富的细节特征,边缘响应值之间的差值绝对值越大,往往表示两个方向之间的细节特征越突出,更突出的细节特征往往具有更强的判别力.此外,基于 Kirsch 3×3 掩模约束机制的分析,DLDP 为高分辨率的 Kirsch 掩模设计合理的权值.实验结果显示:DLDP 在光照、表情、姿态和遮挡方面获取了较好的结果;尤其针对遮挡情况,DLDP 表现得更为突出.

致谢 在此,我们向对本文的工作给予支持和建议的同行,尤其是武汉大学软件工程国家重点实验室丁立新教授、余旌胡教授领导的讨论班上的老师和同学表示感谢.

© 中国科学院软件研究所 http://www.jos.org.cn

References:

[1] Zhao W, Chellappa R, Phillips PJ, Rosenfeld A. Face recognition: A literature survey. ACM Computing Surveys, 2003,34(4): 399−485. [doi: 10.1145/954339.954342]

[2] Jain AK, Ross A, Prabhakar S. An introduction to biometric recognition. IEEE Trans. on Circuits and Systems for Video Technology, 2004,14(1):4−20. [doi: 10.1109/TCSVT.2003.818349]

[3] Jafri R, Arabnia HR. A survey of face recognition techniques. Journal of Information Processing System, 2009,5(2):41−68. [doi:10.3745/JIPS.2009.5.2.041]

[4] Turk M, Pentland A. Eigenfaces for recognition. Journal of Cognitive Neuroscience, 1991,3(1):71−86. [doi: 10.1162/jocn.1991.3.1.71]

[5] Comon P. Independent component analysis—A new concept. Signal Process, 1994,36(3):287−314. [doi: 10.1016/0165-1684(94)90029-9]

[6] Belhumeur PN, Hespanha JP, Kriegman DJ. Eigenfaces versus Fisherfaces: Recognition using class specific linear projection. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1997,19(7):711−720. [doi: 10.1109/34.598228]

[7] Yang J, Zhang D, Frangi AF, Yang J. Two-Dimensional PCA: A new approach to appearance-based face representation and recognition. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2004,26(1):131−137. [doi: 10.1109/TPAMI.2004.1261097]

[8] Tenenbaum JB, De Silva V, Langford JC. A global geometric framework for nonlinear dimensionality reduction. Science, 2000, 290(5500):2319−2323. [doi: 10.1126/science.290.5500.2319]

[9] Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. Science, 2000,290(5500):2323−2326. [doi: 10.1126/science.290.5500.2323]

[10] He X, Yan S, Hu Y, Niyogi P, Zhang H. Face recognition using laplacianfaces. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005,27(3):328−340. [doi: 10.1109/TPAMI.2005.55]

[11] He X, Cai D, Yan S, Zhang H. Neighborhood preserving embedding. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2005. 1208−1213. [doi: 10.1109/ICCV.2005.167]

[12] Yan SC, Xu D, Zhang BY. Graph embedding and extensions: A general framework for dimensionality reduction. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007,29(1):40−51. [doi: 10.1109/TPAMI.2007.250598]

[13] Li ZK, Ding LX, Wang Y. A robust dimensionality reduction method from laplacian orientations. In: Proc. of the 15th IEEE Int'l Conf. on High Performance Computing and Communications. 2013. 345−351. [doi: 10.1109/HPCC.and.EUC.2013.57]

[14] Li ZK, Ding LX, Wang Y, He JR. Face representation with gradient orientations and euler mapping: Application to face recognition. Int'l Journal of Pattern Recognition and Artificial Intelligence, 2014,28(8):1−21. [doi: 10.1142/S021800141456014X]

[15] Belkin M, Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering. In: Proc. of the NIPS, Vol.14. 2001. 585−591. http://machinelearning.wustl.edu/mlpapers/paper_files/nips02-AA42.pdf

[16] Liu C, Wechsler H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. IEEE Trans. on Image Process, 2002,11(4):467−476. [doi: 10.1109/TIP.2002.999679]

[17] Ahonen T, Hadid A, Pietikainen M. Face description with local binary patterns: Application to face recognition. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2006,28(12):2037−2041. [doi: 10.1109/TPAMI.2006.244]

[18] Tan XY, Triggs B. Enhanced local texture feature sets for face recognition under difficult lighting conditions. IEEE Trans. on Image Processing, 2010,19(6):1635−1650. [doi: 10.1109/TIP.2010.2042645]

[19] Zhang WC, Shan SG, Gao W, Zhang HM. Local gabor binary pattern histogram sequence (lgbphs): A novel non-statistical model for face representation and recognition. In: Proc. of the IEEE Int'l Conf. and Computer Vision. 2005. 786−791. [doi: 10.1109/ICCV.2005.147]

[20] Heikkilä M, Pietikäinen M, Schmid C. Description of interest regions with local binary patterns. Pattern Recognition, 2009,42(3): 425−436. [doi: 10.1016/j.patcog.2008.08.014]

[21] Lei Z, Liao SC, Pietikainen M, Li SZ. Face recognition by exploring information jointly in space, scale and orientation. IEEE Trans. on Image Processing, 2011,20(1):247−256. [doi: 10.1109/TIP.2010.2060207]

© 中国科学院软件研究所 http://www.jos.org.cn

[22] Jabid T, Kabir MH, Chae O. Robust facial expression recognition based on local directional pattern. ETRI Journal, 2010,32(5): 784−794. [doi: 10.4218/etrij.10.1510.0132]

[23] Zhong FJ, Zhang JS. Face recognition with enhanced local directional patterns. Neurocomputing, 2013,119:375−384. [doi: 10.1016/j.neucom.2013.03.020]

[24] Tzimiropoulos G, Zafeiriou S, Pantic M. Subspace learning from image gradient orientations. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2012,34(12):2454−2466. [doi: 10.1109/TPAMI.2012.40]

[25] Zhang TP, Tang YY, Fang B, Shang ZW, Liu XY. Face recognition under varying illumination using gradientfaces. IEEE Trans. on Image Processing, 2009,18(11):2599−2605. [doi: 10.1109/TIP.2009.2028255]

[26] Ojala T, Pietikainen M. 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. [doi: 10.1109/TPAMI.2002.1017623]

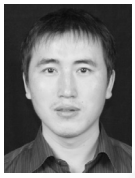
[27] Martinez AM, Benavente R. The AR face database. CVC Technical Report, #24, 1998.

[28] Lee KC, Ho J, Kriegman DJ. Acquiring linear subspaces for face recognition under variable lighting. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005,27(5):684−698. [doi: 10.1109/TPAMI.2005.92]

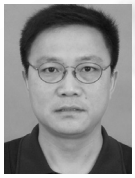
[29] Sim T, Baker S, Bsat M. The CMU pose, illumination, and expression database. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2003,25(12):1615−1618. [doi: 10.1109/TPAMI.2003.1251154]

[30] Cai D, He XF, Han JW. Spectral regression for efficient regularized subspace learning. In: Proc. of the IEEE 11th Int'l Conf. on Computer Vision. 2007. 1−8. [doi: 10.1109/ICCV.2007.4408855]

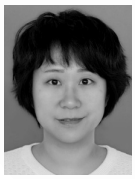
[31] Phillips PJ, Moon H, Rizvi SA, Rauss PJ. The FERET evaluation methodology for face recognition algorithms. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2000,22(10):1090−1104. [doi: 10.1109/34.879790]



李照奎(1976－),男,河南邓州人,博士,副教授,CCF 学生会员,主要研究领域为机器学习,人脸识别,图像处理.

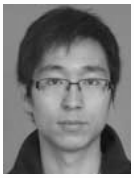


丁立新(1967－),男,博士,教授,博士生导师,CCF 会员,主要研究领域为智能信息处理,云计算.



王岩(1978－),女,讲师,CCF 会员,主要研究领域为机器学习,图像处理.

何进荣(1984－),男,博士,主要研究领域为机器学习,数据降维.



丁国辉(1982－),男,博士,讲师,主要研究领域为机器学习,数据集成,数据库模式匹配.

© 中国科学院软件研究所 http://www.jos.org.cn