

基于图像分解的人脸特征表示^{*}

李照奎^{1,2}, 丁立新¹, 何进荣¹, 胡庆辉¹

¹(软件工程国家重点实验室(武汉大学 计算机学院),湖北 武汉 430072)

²(沈阳航空航天大学 计算机学院,辽宁 沈阳 110136)

通讯作者: 丁立新, E-mail: lxding@whu.edu.cn, http://www.whu.edu.cn

摘要: 提出一种基于图像分解的人脸特征表示方法(FRID),首先通过多方向操作,把一幅图像分解成一系列方向子图像;然后,通过欧拉映射操作,把每幅方向子图像分解成实部和虚部图像,针对每幅实部和虚部图像,分别划分出多个不重叠的局部图像块,通过统计图像块上不同数值的个数生成相应的实部和虚部直方图,一幅图像的所有实部和虚部直方图被串联成一个超级特征向量;最后,利用线性判别分析方法对超级特征向量进行维数约简,以获得每幅图像的低维表示.实验显示该方法在多个人脸数据库上获得了优于时新算法的识别结果,并且表现得更为稳定.

关键词: 图像分解;多方向操作;欧拉映射;人脸识别

中图法分类号: TP391

中文引用格式: 李照奎,丁立新,何进荣,胡庆辉.基于图像分解的人脸特征表示.软件学报,2014,25(9):2102-2118. <http://www.jos.org.cn/1000-9825/4651.htm>

英文引用格式: Li ZK, Ding LX, He JR, Hu QH. Face feature representation based on image decomposition. Ruan Jian Xue Bao/ Journal of Software, 2014, 25(9): 2102-2118 (in Chinese). <http://www.jos.org.cn/1000-9825/4651.htm>

Face Feature Representation Based on Image Decomposition

LI Zhao-Kui^{1,2}, DING Li-Xin¹, HE Jin-Rong¹, HU Qing-Hui¹

¹(State Key Laboratory of Software Engineering (School of Computer, Wuhan University), Wuhan 430072, China)

²(School of Computer, Shenyang Aerospace University, Shenyang 110136, China)

Corresponding author: DING Li-Xin, E-mail: lxding@whu.edu.cn, <http://www.whu.edu.cn>

Abstract: This paper presents a face feature representation method based on image decomposition (FRID). FRID first decomposes an image into a series of orientation sub-images by executing multiple orientations operator. Then, each orientation sub-image is decomposed into a real part image and an imaginary part image by applying Euler mapping operator. For each real and imaginary part image, FRID divides them into multiple non-overlapping local blocks. The real and imaginary part histograms are calculated by accumulating the number of different values of image blocks respectively. All the real and imaginary part histograms of an image are concatenated into a super-vector. Finally, the dimensionality of the super-vector is reduced by linear discriminant analysis to yield a low-dimensional, compact, and discriminative representation. Experimental results show that FRID achieves better results in comparison with state-of-the-art methods, and is the most stable method.

Key words: image decomposition; multiple orientations operator; Euler mapping; face recognition

人脸识别被广泛应用在信息安全和访问控制、法律实施、人机接口以及更普遍的图像理解上^[1],如何提取人脸特征是一个关键步骤.一般来说,人脸特征可以分为全局特征和局部特征两大类:全局特征是指其特征向量的每一维都包含了人脸图像上所有部分(甚至所有像素)的信息,因此反映的是人脸的整体属性;与全局特征不

* 基金项目: 国家自然科学基金(60975050, 60902053, 61170185); 广东省省部产学研结合专项资金(2011B090400477); 珠海市产学研合作专项资金(2011A050101005, 2012D0501990016); 珠海市重点实验室科技攻关项目(2012D0501990026)

收稿时间: 2014-04-06; 修改时间: 2014-05-14; 定稿时间: 2014-06-09

同,局部特征的每一维都只对应人脸图像上的一个局部区域,因此侧重于提取人脸的细节特征.在人脸识别的研究中,基于子空间方法是一类主要的全局特征,该类方法主要把人脸图像投影到低维子空间上,主成分分析(principal component analysis,简称 PCA)^[2]和线性判别分析(linear discriminant analysis,简称 LDA)^[3]作为两个著名的线性子空间学习方法,已经被广泛应用在模式识别和计算机视觉上.Yan 等人^[4]从图嵌入角度重新解释了子空间学习方法,从而把各种子空间方法,如 ISOMAP^[5],LLE^[6], LPP^[7],NPE^[8],MFA^[4]等纳入到统一的图嵌入框架.进而,为了处理非线性的人脸特征空间,各种基于核的方法把线性特征表示方法(如 PCA 和 LDA)扩展到任意的希尔伯特空间进行^[9,10].这些方法的主要思想是:通过一个隐式的非线性映射,把输入数据映射到高维的希尔伯特空间,然后在该高维特征空间上执行线性特征提取方法,所有的计算都通过特征空间的内积进行.一般的线性特征算法都有其相对应的核算法.

在面对表情、姿态、光照、遮挡时,基于全局特征识别算法的性能明显下降.近年来,局部特征被认为对人脸的光照、表情和遮挡等变化不敏感,因此被越来越多地用于人脸特征表示.已有的局部特征主要分为两类:(1)稀疏局部特征;(2)稠密局部特征.稀疏局部特征首先探测给定图像的关键点,然后对局部块进行采样并形成相应的不变特征.尺度不变特征变换(scale invariant feature transform,简称 SIFT)^[11]和梯度方向直方图(histogram of oriented gradients,简称 HOG)^[12]是两种典型的稀疏局部特征.稠密局部特征通常直接基于像素把人脸图像划分多个局部块,在每块中,特征分别被提取,然后利用信息融合策略把多个块的特征融合成一个单个特征去识别,或者融合每一块的识别结果.比较常用的稠密局部特征包括局部二值模式(local binary pattern,简称 LBP)^[13]、Gabor 小波^[14,15]等.Gabor 小波主要获取相应的空间频率(尺度)、空间位置和方向的局部结构,实验证实,Gabor 小波特征对光照和表情变化比较稳健.LBP 描述中心点与四周邻居点的变化情况,在某种程度上对于光照变化比较稳健.许多 LBP 的改进算法已经取得成功^[16,17],随后,通过组合 LBP 和 Gabor 的一些局部特征也已被提出.相比于单个特征,这些组合特征获得了更好的识别效果.文献[18]提出在 Gabor 幅值上进行 LBP 特征提取.文献[19]提出在 Gabor 相位上进行 LBP 特征提取.在文献[20]中,不仅仅通过图像空间,而且通过多尺度多方向的 Gabor 滤波去进行 LBP 特征提取.文献[21]提出融合 Gabor 幅值和 Gabor 相位进行 LBP 特征提取.另外也有通过组合局部和全局特征来获取新的局部特征并获得较好的识别效果.在 2008 年和 2012 年,文献[22,23]分别提出利用傅里叶频谱的相位信息进行 LBP 特征提取(LPQ).2011 年,文献[24]提出利用傅里叶频谱的幅度信息和相位信息进行 LBP 特征提取(LFD).2013 年,文献[25]在 LPQ 的基础上提出了多尺度的 LPQ,并使用核级融合策略对多个 LBP 特征进行融合.2002 年,文献[15]提出 Gabor 特征和 FLDA 特征的组合进行人脸识别.2009 年,文献[26]提出 Gabor 特征和傅里叶变换后的低频特征集合方法进行人脸识别.近来,基于稀疏表示的分类方法(SRC)^[27]被应用到人脸识别后,越来越多的研究人员把研究兴趣转移到稀疏编码的研究上,比较流行的方法有基于协作表示的分类方法(CRC)^[28]、基于宽松的协作表示分类方法(RCR)^[29].

在 2013 年,Qian 等人^[30]提出一种基于局部结构信息的特征表示方法,该方法利用回归对图像进行分解以获得图像局部结构信息,然后,结合 LDA 进行维数约简.受此启发,本文提出一种新的基于图像分解的人脸特征表示方法 FRID.为了获取更加丰富的局部结构信息,本文引入多方向操作(包括多交叉梯度方向和多拉普拉斯方向),通过多方向操作,一幅输入图像被分解成多幅不同的方向子图像;此外,通过引入鲁棒的余弦相关性度量方式^[31],进而得出一种明确的非线性欧拉映射方法,通过执行该欧拉映射,每幅方向子图像被分解成两幅不同的图像:实部图像和虚部图像;针对实部图像和虚部图像,分别划分出多个不重叠的局部图像块,通过统计子图像块上不同数值的个数,生成相应的实部直方图和虚部直方图,所有的实部直方图和虚部直方图被串联成一个新的超级特征向量;最后,通过 LDA 对超级特征向量进行维数约简,以获得每幅输入图像的低维表示方法.FRID 方法的整体流程如图 1 所示.

本文第 1 节介绍多方向图像分解.第 2 节提出相应的欧拉分解方法.第 3 节设计相应的直方图表示方法.第 4 节描述相应的维数约简方法.实验结果与分析在第 5 节.最后,在第 6 节进行总结.

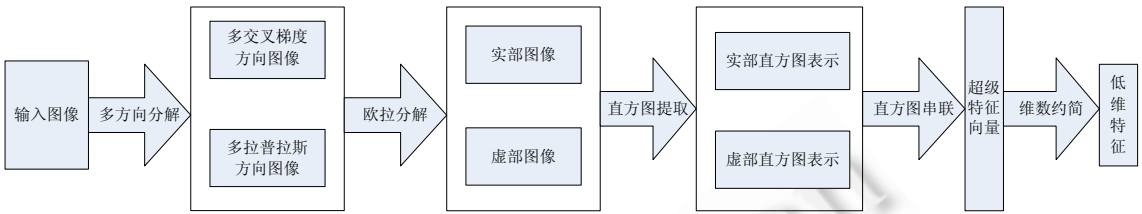


Fig.1 Overview of the FRID method

图 1 FRID 的整体流程

1 多方向图像分解

1.1 多交叉梯度方向图像分解

梯度算子是一阶微分算子,主要目的是增强图像轮廓,突出图像细节.对于输入图像 $F(x,y)$,它的梯度值是一个向量:

$$\nabla F(x,y) = \begin{pmatrix} G_x \\ G_y \end{pmatrix} = \begin{pmatrix} \partial F / \partial x \\ \partial F / \partial y \end{pmatrix} \quad (1)$$

该向量的幅值被表示为

$$mag(F) = (G_x^2 + G_y^2)^{1/2} \quad (2)$$

该向量的方向角表示为

$$\alpha(x,y) = \arctan\left(\frac{G_y}{G_x}\right) \quad (3)$$

假定在图像 $F(x,y)$ 中有 $M \times N$ 个像素,把图像中第 i 个像素作为中心,然后用 city-block 距离在一个半径为 R 的正方形中选择它的 t 个邻居像素,这些邻居像素形成一个方形的对称邻居集 $\Omega_i^t = \{f_j | j=1, \dots, t\}$. 图 2 列出了某个中心像素的不同邻居集.针对图像中的每个像素,通常通过不同的模板来计算它的梯度信息,图 3 展示了半径为 1 的图像模板 T ,在模板 T 中, f_c 表示中心位置像素的强度值, $f_j(j=1,2, \dots, 8)$ 表示 f_c 的第 j 个邻居.为了获取更丰富的局部结构信息,在半径为 1 的图像模板中引入 8 种交叉梯度方向操作,具体交叉操作模板如图 4 所示.图 4(a)~图 4(h) 分别表示 8 种不同的交叉梯度方向操作,这些操作共同的特点是, G_y 和 G_x 分别交叉计算,图 4(a)~图 4(h) 所对应的 8 种交叉梯度方向计算见公式(4)~公式(11).

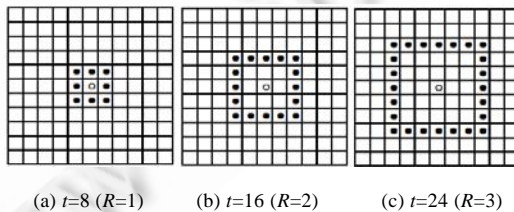


Fig.2 Squarely symmetric neighbor sets for radius R

图 2 不同半径下的方形对称邻居集

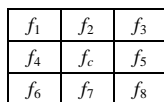


Fig.3 Image template ($R=1$)

图 3 图像模板($R=1$)

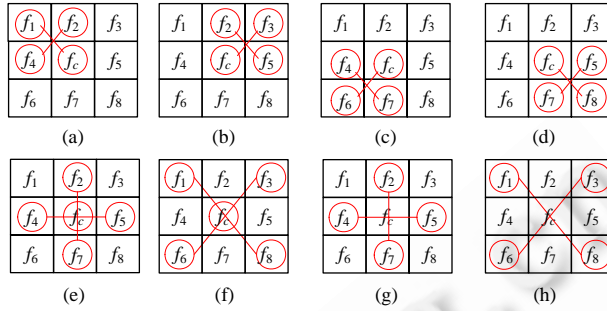


Fig.4 Cross-Gradient orientation operator

图 4 交叉梯度方向操作

$$G_y^1 = f_1 - f_c, G_x^1 = f_2 - f_4, \alpha^1(x, y) = \arctan\left(\frac{G_y^1}{G_x^1}\right) \quad (4)$$

$$G_y^2 = f_3 - f_c, G_x^2 = f_2 - f_5, \alpha^2(x, y) = \arctan\left(\frac{G_y^2}{G_x^2}\right) \quad (5)$$

$$G_y^3 = f_6 - f_c, G_x^3 = f_7 - f_4, \alpha^3(x, y) = \arctan\left(\frac{G_y^3}{G_x^3}\right) \quad (6)$$

$$G_y^4 = f_8 - f_c, G_x^4 = f_7 - f_5, \alpha^4(x, y) = \arctan\left(\frac{G_y^4}{G_x^4}\right) \quad (7)$$

$$G_y^5 = f_2 + f_7 - 2f_c, G_x^5 = f_4 + f_5 - 2f_c, \alpha^5(x, y) = \arctan\left(\frac{G_y^5}{G_x^5}\right) \quad (8)$$

$$G_y^6 = f_1 + f_8 - 2f_c, G_x^6 = f_3 + f_6 - 2f_c, \alpha^6(x, y) = \arctan\left(\frac{G_y^6}{G_x^6}\right) \quad (9)$$

$$G_y^7 = f_1 - f_8, G_x^7 = f_3 - f_6, \alpha^7(x, y) = \arctan\left(\frac{G_y^7}{G_x^7}\right) \quad (10)$$

$$G_y^8 = f_2 - f_7, G_x^8 = f_4 - f_5, \alpha^8(x, y) = \arctan\left(\frac{G_y^8}{G_x^8}\right) \quad (11)$$

针对多交叉梯度方向算子设计的 8 种模板,目的是为了获取不同区域的边缘信息:

- 对于模板(a),计算方法对应公式(4),主要为了获取参考点 f_c 左上区域(f_1, f_2, f_4, f_c)的交叉边缘信息($f_1 \leftrightarrow f_c$ 走向的边缘或 $f_2 \leftrightarrow f_4$ 走向的边缘),如果 $\alpha^1=0$,则代表在左上区域拥有 $f_2 \leftrightarrow f_4$ 走向的边缘信息;
- 同理,通过模板(b)获取参考点 f_c 右上区域(f_2, f_3, f_c, f_5)的交叉边缘信息;
- 通过模板(c)获取参考点 f_c 左下区域(f_4, f_c, f_6, f_7)的交叉边缘信息;
- 通过模板(d)获取参考点 f_c 右下区域(f_c, f_5, f_7, f_8)的交叉边缘信息;
- 而对于模板(g)和模板(h),都是为了获取交叉穿过参考点 f_c 的边缘信息,其中,模板(g)获取的是倾斜交叉,模板(h)获取的是垂直交叉;
- 至于模板(e)和模板(f),分别是模板(g)和模板(h)更细化的设计,相对于模板(g)和模板(h),模板(e)和模板(f)获取的边缘信息更细节化。

针对图像 $F(x,y)$ 的任意一个像素,首先根据公式(4)~公式(11)分别计算它的 8 个交叉梯度方向 $\alpha^s(x,y)$ ($s=1, \dots, 8$),最终,一幅输入图像 $F(x,y)$ 将被分解为 8 幅不同的交叉梯度方向图像 $\Phi^s(x,y)$ ($s=1, \dots, 8$),图 5 展示了多交叉梯度方向图像的计算实例。

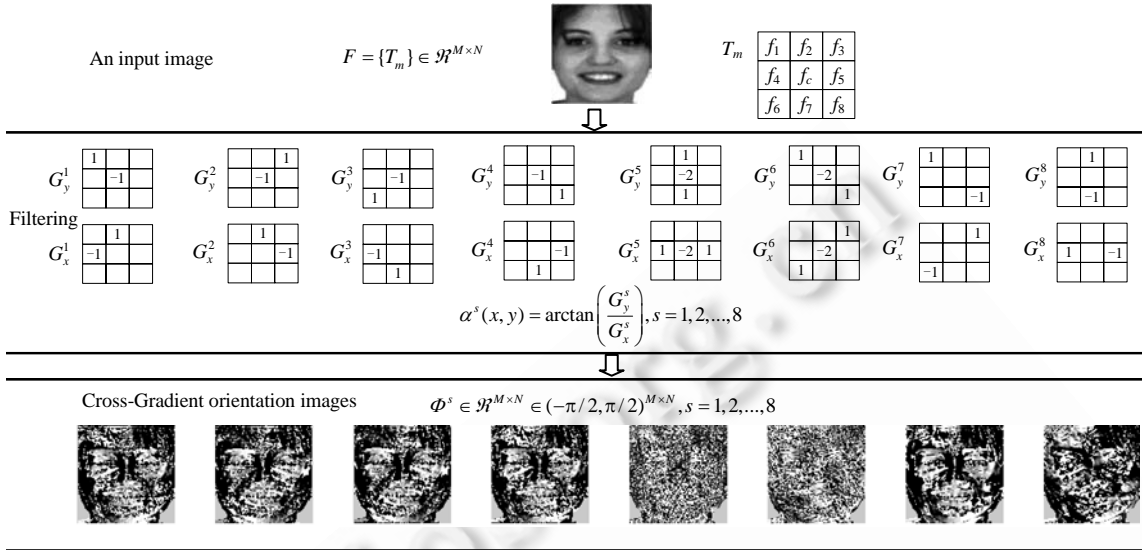


Fig.5 Illustration of the computation of multiple cross-gradient orientation images
图5 多交叉梯度方向图像的计算实例

1.2 多拉普拉斯方向图像分解

拉普拉斯算子是二阶微分算子,具有各向同性和位移不变性,从而满足不同走向图像边缘的锐化要求.对于连续图像 $F(x,y)$,其拉普拉斯算子为

$$\nabla^2 F = \frac{\partial^2 F(x, y)}{\partial x^2} + \frac{\partial^2 F(x, y)}{\partial y^2} \tag{12}$$

对于数字图像 $F(x,y)$,其拉普拉斯算子为

$$\nabla^2 F = \Delta_x^2 F(x, y) + \Delta_y^2 F(x, y) \tag{13}$$

其中,

$$\Delta_x^2 F(x, y) = \Delta_x[\Delta_x F(x, y)] = F(x+1, y) + F(x-1, y) - 2F(x, y) \tag{14}$$

同理,

$$\Delta_y^2 F(x, y) = F(x, y+1) + F(x, y-1) - 2F(x, y) \tag{15}$$

公式(13)中的二维拉普拉斯算子实现可由两个分量相加得到:

$$\nabla^2 F = [F(x+1, y) + F(x-1, y) + F(x, y+1) + F(x, y-1)] - 4F(x, y) \tag{16}$$

公式(16)给出了以 90° 旋转的各向同性的结果.对角线方向也可以加入到离散拉普拉斯变换的定义中,由于每个对角线方向上的项还包含一个 $-2f(x,y)$,所以总共减去 $8f(x,y)$.以图3中半径为1的图像模板为例,中心像素 f_c 的两个拉普拉斯操作可以计算如下:

$$\nabla^2 F = (f_2 + f_4 + f_5 + f_7 - 4f_c) \tag{17}$$

$$\nabla^2 F = \left(\sum_{k=1}^8 f_k - 8f_c \right) \tag{18}$$

通常假定图像中相邻像素的强度值比较接近,为了平滑变化过大的像素差值,引入拉普拉斯方向操作,具体定义如下:

$$\beta^1(x, y) = \arctan[(f_2 + f_4 + f_5 + f_7 - 4f_c)/4] \tag{19}$$

$$\beta^2(x, y) = \arctan\left[\left(\sum_{k=1}^8 f_k - 8f_c\right)/8\right] \tag{20}$$

拉普拉斯方向操作如图 6 所示,其中,图 6(a)、图 6(b)分别对应公式(19)、公式(20).针对图像 $F(x,y)$ 的任意一个像素,首先根据公式(19)、公式(20)分别计算它的 2 个拉普拉斯方向 $\beta^t(x,y)(t=1,2)$,最终,一副输入图像 $F(x,y)$ 将被分解为 2 幅不同的拉普拉斯方向图像 $I^t(x,y)(t=1,2)$.图 7 展示了两个拉普拉斯方向图像的计算实例.对于一幅输入图像 $F(x,y)$,通过执行交叉梯度方向和拉普拉斯方向操作,分别得到 8 幅交叉梯度方向图像和 2 幅拉普拉斯方向图像.在方向值的计算中,通过反正切函数来平滑变化过快的像素强度差,输入图像 $F(x,y)$ 中的像素值取值范围在 $[0,255]$,经过分解后的方向子图像的每个像素值的取值被限制在 $(-\pi/2,\pi/2)$.多交叉梯度方向和多拉普拉斯方向的计算,从多个角度加强了图像的边缘信息和细节信息.

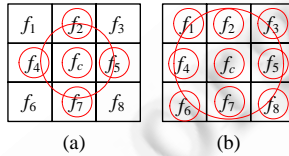


Fig.6 Laplacian orientation operator

图 6 拉普拉斯方向操作

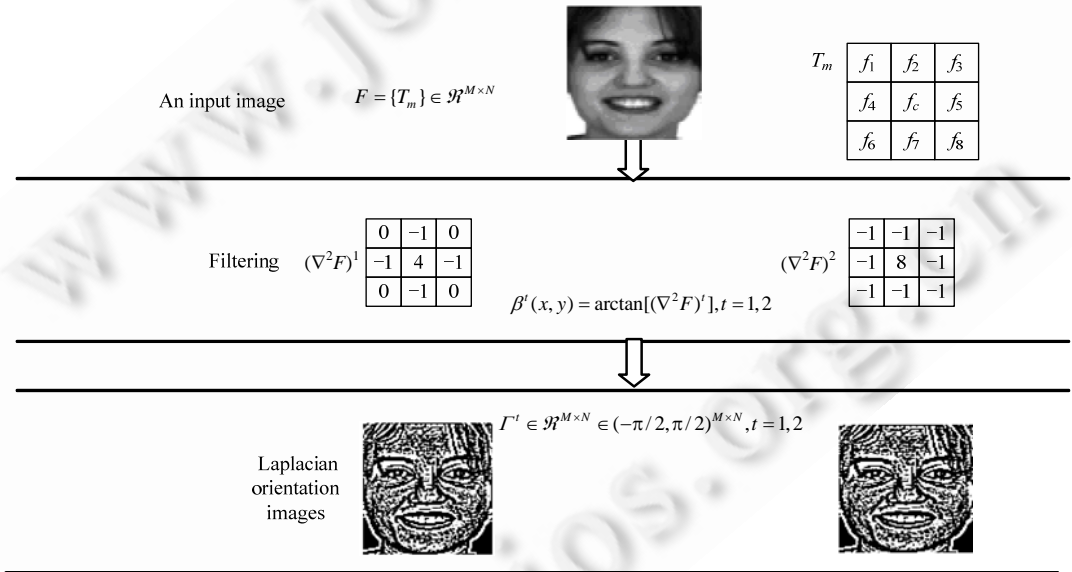


Fig.7 Illustration of the computation of multiple Laplacian orientation images.

图 7 拉普拉斯方向图像的计算实例

2 欧拉图像分解

对于给定的输入图像 $F \in \mathcal{R}^{M \times N}$,根据第 1.1 节和第 1.2 节可以计算相应的交叉梯度方向图像 $\Phi^s(x,y)(s=1,\dots,8)$ 和拉普拉斯方向图像 $I^t(x,y)(t=1,2)$.为方便起见,统一用 $O \in \mathcal{R}^{M \times N}$ 表示两类方向图像,用 b_j, b_k 分别表示图像 O_j, O_k 对应的列向量,则对图像 O_j, O_k 间的相关性计算可用如下鲁棒的余弦函数计算^[31]:

$$d(b_j, b_k) = \sum_{q=1}^p \{1 - \cos(b_j^{(q)} - b_k^{(q)})\} \quad (21)$$

这里, $p=M \times N$, $b_j^{(q)}$ 表示 b_j 的第 q 维.文献[33]显示:公式(21)等价于 Andrews 的 M -估计,是一种更鲁棒的图像相关性度量方式.基于公式(21)进行如下推导:

$$\begin{aligned}
d(b_j, b_k) &= \sum_{q=1}^p \{1 - \cos(b_j^{(q)} - b_k^{(q)})\} \\
&= \frac{1}{2} \sum_{q=1}^p \{2 - 2(\cos b_j^{(q)} \cos b_k^{(q)} + \sin b_j^{(q)} \sin b_k^{(q)})\} \\
&= \frac{1}{2} \sum_{q=1}^p \{(\cos b_j^{(q)} - \cos b_k^{(q)})^2 + (\sin b_j^{(q)} - \sin b_k^{(q)})^2\} \\
&= \frac{1}{2} \|(\cos b_j + i \sin b_j) - (\cos b_k + i \sin b_k)\|_F^2 \\
&= \left\| \frac{1}{\sqrt{2}} e^{ib_j} - \frac{1}{\sqrt{2}} e^{ib_k} \right\|_F^2
\end{aligned} \tag{22}$$

这里设:

$$z_j = \frac{1}{\sqrt{2}} \left[e^{ib_j^{(1)}}, e^{ib_j^{(2)}}, \dots, e^{ib_j^{(p)}} \right]^T = \frac{1}{\sqrt{2}} e^{ib_j} \tag{23}$$

故,公式(22)可重新写为

$$d(b_j, b_k) = \|z_j - z_k\|_F^2 \tag{24}$$

公式(24)显示,把方向图像之间的余弦相关距离转变成了希尔伯特空间的欧式距离.很明显,对于一幅图像,可以通过公式(23)进行重新表示.由于公式(21)的鲁棒性,公式(23)实际上是一种新的鲁棒的表示方法.这里,公式(23)可以变换为

$$z_j = \frac{1}{\sqrt{2}} e^{ib_j} = \frac{1}{\sqrt{2}} (\cos b_j + i \sin b_j) \tag{25}$$

根据公式(25),一幅图像被分解成一幅实部图像和一幅虚部图像,此分解被称为欧拉分解,分解结果如下:

$$\begin{aligned}
\xi_j^{real} &= \text{real}(z_j) = \frac{1}{\sqrt{2}} \cos(b_j) \\
\xi_j^{imag} &= \text{imag}(z_j) = \frac{1}{\sqrt{2}} \sin(b_j)
\end{aligned} \tag{26}$$

这里, $\xi_j^{real} \in (0, 1/\sqrt{2}]^p$, $\xi_j^{imag} \in (-1/\sqrt{2}, 1/\sqrt{2})^p$, 实部图像对应余弦函数,虚部图像对应正弦函数.

第 1 节的多方向操作通常捕获了更多的图像边缘信息,方向图像实际是高频图像.方向图像的计算把每个像素对应的方向值限定在一个更小的区间 $(-\pi/2, \pi/2)$ 中.图 8 展示了在该区间中,实部和虚部函数对方向值的影响.给定方向图像中的一个方向值 $b_j^{(q)}, b_j^{(q)}$ 是一个有符号值,用 $|b_j^{(q)}|$ 表示方向值的幅度值,正负值分别表示方向值的正向和负向.在图 8 中,虚部函数的变化趋势近似于直线函数,因此在区间 $(-\pi/2, \pi/2)$ 中,对于不同的方向值,虚部函数实际上起到了保持的作用.也就是说,虚部函数保留了方向值的方向性(正负方向保持不变),零值左右的方向值完全保持不变(包括幅度值和方向).随着方向值变大(向 $\pi/2$ 趋近)或变小(向 $-\pi/2$ 趋近),虚部函数起到了抑制过快变化的作用.因此可以认为:虚部函数一方面保留了方向图像的主要信息,另一方面又起到了对方向图像的平滑作用.由于方向图像是高频图像,因此虚部图像是一种具有一定平滑程度的高频图像.针对实部函数,从图 8 可以看出:所有方向值的方向性都被转为正向,零值左右方向值的幅度值被实部函数放大;随着方向值变大(向 $\pi/2$ 趋近)或变小(向 $-\pi/2$ 趋近),实部函数压低了方向值的幅度值.与虚部图像相比,实部图像中的所有值都变为正值,因此相比于高频的虚部图像,实部图像显然是一种低频图像.图 9 展示了一幅原始的微笑图像:多方向分解图像、欧拉分解图像(实部图像和虚部图像),以及这些图像相应的傅里叶频谱.从图 9 可以看出:虚部图像的频谱和方向图像的频谱近似,因此虚部图像属高频图像.而从实部图像的频谱可以明显看出,实部图像属于低频图像.针对方向图像进行的欧拉分解,实际上综合考虑了方向图像的高频信息和低频信息,因此具有更好的判别性.

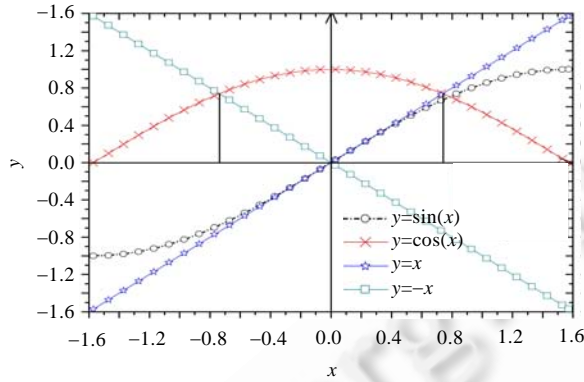


Fig.8 Effect of the real part and the imaginary part

图 8 实部和虚部的作用

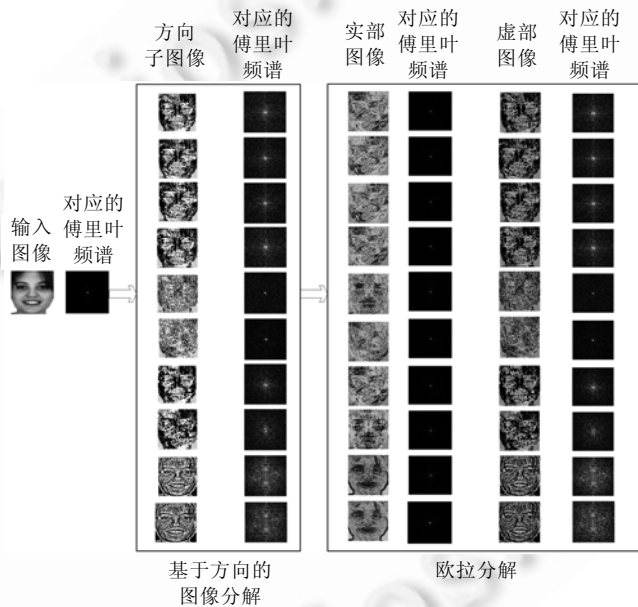


Fig.9 An original smile image , the orientation images, the imaginary part images, the real part images , and the corresponding Fourier spectrums of these images

图 9 原始微笑图像、方向图像、虚部图像和实部图像,以及这些图像相应的傅里叶频谱

3 直方图表示

本节主要针对欧拉分解后的实部图像 E_{real} 和虚部图像 E_{imag} 设计空间直方图编码局部图像块的统计信息, 这里把 $E_{real} \in (0, 1/\sqrt{2}]$ 和 $E_{imag} \in (-1/\sqrt{2}, 1/\sqrt{2})$ 分开表示, 主要因为两类图像的像素取值范围不一致. 为了获得更有效的直方图表示, 分别把实部和虚部图像划分成 L 个非重叠的局部图像块 $E_{real}=[b^1, \dots, b^j, \dots, b^L], b^j \in \mathcal{R}^{w \times w}$ 和 $E_{imag}=[c^1, \dots, c^j, \dots, c^L], c^j \in \mathcal{R}^{w \times w}$, 其中, $w=3, 5$. 由于实部图像和虚部图像的直方图表示方法类似, 下面仅对实部图像的直方图表示进行详细说明.

对于实部图像的每个局部块 $b^j(j=1, 2, \dots, L)$, 计算其相应的子直方图 $H_{real}^j=[A_k^j, \dots, A_k^j, \dots, A_r^j](j=1, 2, \dots, L), r$ 表示连续的取值区间个数(例如 $r=4$), A_k^j 通过累加第 j 个子块中落入第 k 个区间的像素个数来计算. 子直方图 H_{real}^j

的横坐标表示 r 个相等的连续取值区间,纵坐标表示落入每个区间中的像素个数.对于一幅实部图像,串联所有的子直方图 H_{real}^j 以形成一个更具判别性的直方图 $H_{real} = [(H_{real}^1)^T \dots (H_{real}^j)^T \dots ((H_{real}^L)^T)^T]$,用 h_{real} 表示 D 维 ($D=L \times r$)实部列向量.图 10 展示了一幅方向图像的实部直方图和实部列向量的表示过程,同理可以得到一幅方向图像的虚部直方图和虚部列向量.

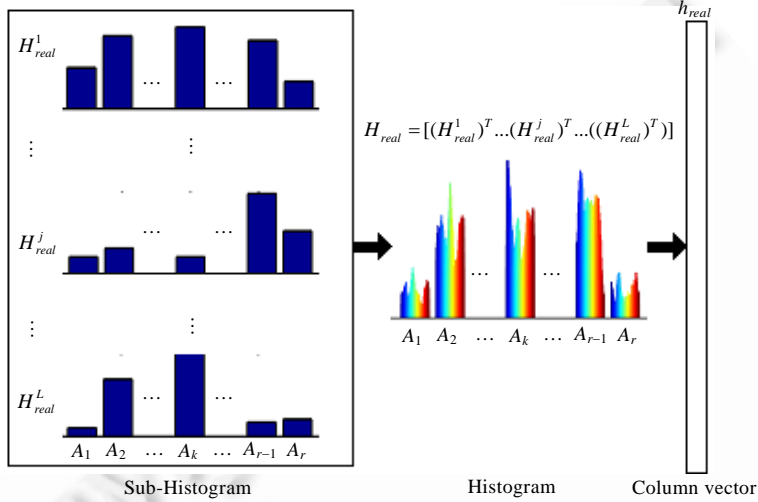


Fig.10 Illustration of the computation of the real histogram of an orientation image

图 10 一幅方向图像的实部直方图计算实例

为了综合考虑实部和虚部图像,下面着重描述实部直方图和虚部直方图的合并过程.假定有 n 个输入图像 $\{F_i\}$,在半径为 1 的局部邻域内,通过多方向分解,每个输入图像被分解成 10 个方向图像 $\{O_i^s\}$.每个方向图像 O_i^s ,又被欧拉方法解成实部图像 $E_{i,real}^s$ 和虚部图像 $E_{i,imag}^s$.对于实部图像,可以获得实部直方图 $H_{i,real}^s$ 和对应的列向量 $h_{i,real}^s$ (如图 10 所示),同理可以获得虚部直方图 $H_{i,imag}^s$ 和对应的列向量 $h_{i,imag}^s$.因此,可以得到两个直方图列向量矩阵 $M_{real} = [h_{1,real}^s \dots h_{n,real}^s] \in \mathcal{R}^{D \times n}$ 和 $M_{imag} = [h_{1,imag}^s \dots h_{n,imag}^s] \in \mathcal{R}^{D \times n}$,每个直方图列向量通过 z-score 方法进行标准化.分别用 $\delta_{i,real}^s$ 和 $\delta_{i,imag}^s$ 表示 $h_{i,real}^s$ 和 $h_{i,imag}^s$ 的标准化向量, $\delta_{i,real}^s$ 的第 q 维用如下方法计算:

$$\delta_{i,real}^{s,(q)} = (h_{i,real}^{s,(q)} - \mu_{real}^{s,(q)}) / \sigma_{real}^{s,(q)}, i = 1, \dots, n, q = 1, \dots, D \tag{27}$$

$$\mu_{real}^{s,(q)} = \frac{1}{n} \sum_{i=1}^n h_{i,real}^{s,(q)} \tag{28}$$

$$\sigma_{real}^{s,(q)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (h_{i,real}^{s,(q)} - \mu_{real}^{s,(q)})^2} \tag{29}$$

在以上计算中, $h_{i,real}^{s,(q)}$, $\delta_{i,real}^{s,(q)}$ 分别是 $h_{i,real}^s$ 和 $\delta_{i,real}^s$ 的第 q 维.可以根据公式(27)~公式(29)计算 $\delta_{i,real}^s$ 的所有维.同理可以得到虚部标准化列向量的所有维.对于一幅图像 F_i ,让 $\delta_{i,real}^s$ ($i=1, \dots, n; s=1, \dots, 10$)表示其所对应的实部直方图的标准列向量, $\delta_{i,imag}^s$ ($i=1, \dots, n; s=1, \dots, 10$)表示其所对应的虚部直方图的标准列向量.接下来,把所有的 $\delta_{i,real}^s$ 和 $\delta_{i,imag}^s$ 串联成一个新的超级特征向量 $\chi_i = [(\delta_{i,real}^1})^T \dots (\delta_{i,real}^{10})^T (\delta_{i,imag}^1})^T \dots (\delta_{i,imag}^{10})^T]^T \in \mathcal{R}^P$ ($P=20 \times D$),因此对于一幅输入图像,通过图像分解、直方图表示,最终得到一个更有判别性的超级特征向量.

4 维数约简

由于新的超级特征向量维数较高,接下来用 LDA 进行维数约简.LDA 的基本思想是:将高维的模式样本投

影到最佳鉴别矢量空间,以达到抽取分类信息和压缩特征空间维数的效果,投影后保证模式样本在新的子空间有最大的类间距离和最小的类内距离,即模式在该空间中有最佳的可分离性.因此,它是一种有效的特征抽取方法.使用这种方法,能够使投影后模式样本的类间散布矩阵 S_b 最大,并且同时类内散布矩阵 S_w 最小.因此,它能够保证投影后模式样本在新的空间中有最小的类内距离和最大的类间距离,即,模式在该空间中有最佳的可分离性.

S_b 和 S_w 分别定义如下:

$$S_b = \sum_{c=1}^C l^c (m^c - m)(m^c - m)^T \quad (30)$$

$$S_w = \sum_{c=1}^C \sum_{i=1}^{l^c} (\chi_{i,c} - m^c)(\chi_{i,c} - m^c)^T \quad (31)$$

这里, C 代表总共的类别个数, $\chi_{i,c}$ 是第 c 类中的第 i 张脸, m^c 是第 c 类中所有样本的均值脸, l^c 是第 c 类中的样本个数, m 是所有类中样本的均值. LDA 的子空间通过一系列列向量 Q 扩张而成,这里, Q 满足如下式子:

$$Q = \arg \max \left| \frac{Q^T S_b Q}{Q^T S_w Q} \right| \quad (32)$$

因此, Q 能够通过求解 $S_w^{-1} S_b$ 的特征向量来获得, Q 映射原始的 P 维空间到一个最终的 d 维特征空间.为了避免 S_w 的奇异性,采用如下的正则化方法:

$$(S_w + \varepsilon I)^{-1} S_b Q = Q A \quad (33)$$

这里, ε 是一个小的正则化常量(取 10^{-3} 以下), I 是一个单位矩阵.

对于一个超级特征向量 χ , 最终通过 $\varphi = Q^T \chi$ 投影到一个优化的判别空间, φ 是 χ 的低维特征向量.当进行图像分类时,基于余弦的最近邻分类方法(1-NN)被采用,具体的余弦度量方法见公式(34):

$$d(\varphi_j, \varphi_q) = 1 - \frac{\varphi_j^T \varphi_q}{\sqrt{(\varphi_j^T \varphi_j)(\varphi_q^T \varphi_q)}} \quad (34)$$

这里, φ_j 和 φ_q 分别表示两个不同的低维特征向量.

本文算法的时间复杂度主要包含多方向图像分解、欧拉图像分解、直方图表示和特征分解等 4 个方面,算法运行时间与训练样本个数 n 、维数 p 等有关.多方向图像分解的时间复杂度为 $O(pn)$,欧拉图像分解的时间复杂度为 $O(1)$,直方图表示的时间复杂度为 $O(n)$,特征分解的时间复杂度为 $O(p^3)$.所以,整个算法的时间复杂度为 $O(pn+n+p^3)$.

5 实验结果与分析

本节将在 3 个公开的人脸数据库 AR^[32], Extended Yale B^[33] 和 CMU PIE^[34] 上评估 FRID 算法的效果,并与时新的算法进行比较.图 11 展示了 AR 数据库的图像样例.图 12 展示了 Extended Yale B 数据库的图像样例.图 13 展示了 CMU PIE 数据库的图像样例.在下面的实验中,将设置相关方法的参数.针对提出的 FRID 方法,主要包括维数约简阶段的正则化参数 ε 、多方向计算中的邻居半径 R 、直方图表示中图像分块尺寸 w 和连续的取值区间个数 r .在实验中: $\varepsilon=0.001$;不同的数据库中, R, w 和 r 取值不尽相同.

- 在 AR 中参数设置为 $R=1, w=3$ 和 $r=3$;
- 在 Yale B 参数设置为 $R=2, w=3$ 和 $r=3$;
- 在 PIE 中参数设置为 $R=2, w=3$ 和 $r=3$.

此外,所比较的算法的优化参数设置如下:

- 对于 LBP 特征,在 Extended Yale B 和 PIE 上图像块设置为 8×8 ,在 AR 上图像块设置为 16×15 ,每块对应一个 59 维的直方图;
- 对于 TPLBP 特征, $r_1=2, S=8$ 和 $\omega=5$;
- 对于 LTP 和 LTP/DT 特征,编码的极值设置为 0.02,DT 距离的参数 α 设置为 1,截断 DT 距离的极值设置

为 6,在 AR 上的图像块设置为 16×15 ,其他数据库的图像块设置为 8×8 ;

- 对于 ID-LARK 和 LARK,在 Yale B 上的局部窗体尺寸设置为 3×3 ,其他数据库的设置为 7×7 ;
- 对于比较的 Gabor 特征,每张脸使用 40 个 Gabor 滤波器,包括 5 个尺度和 8 个方向,关于下采样因子在 PIE 中设置为 4,其他数据库中设置为 8;
- 对于 IDLS 特征,岭回归的正则化参数 $\delta=0.01$,下采样因子 $\lambda=2$,邻居半径取 2,宏像素尺寸设置为 3.



Fig.11 Face images of the same subject taken from the AR database

图 11 来自于 AR 的同一个人的多张图像



Fig.12 Partial face images of the same subject taken from the Extended Yale B database

图 12 来自于 Extended Yale B 的同一个人的多张图像



Fig.13 Sample images of a person under different illuminations and five poses in the CMU PIE database

图 13 来自于 CMU PIE 的同一个人的多张图像

5.1 AR人脸数据库

AR 人脸库包括 4 000 多张正面人脸图像,包括 126 个人(70 个男性、56 个女性).每人有 26 张图像,被分作两部分:第 1 部分每个人有 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.120 个人(65 个男性、55 个女性)的 14 张图像(每部分的 1~7)用在本实验中,每张图像的人脸部分被裁剪成 100×90 .在本实验中,第 1 部分的 7 张人脸图像(1~7)被用作训练,第 2 部分的 7 张人脸图像(1~7)被用作测试.然后,PCA,LDA,LARK,ID-LARK,LBP+Chi2,LTP+Chi2,LTP/DT,TPLBP+Chi2,Gabor+LDA,IDLS-Distance,IDLS 与本文提出的 FRID 分别被用来进行特征提取.在 LDA 的 PCA 阶段,主成分的个数被设置为 200.NN 分类器被用来分类,所有方法的识别结果被列举在表 1 中.

Table 1 Recognition rates of different methods on the ar database

表 1 在 AR 人脸库中不同识别方法的识别率

Methods	Recognition rate (%)
PCA	60.7
LDA	78.1
LARK	82.5
ID-LARK	91.9
LBP+Chi2	91.3
LTP+Chi2	90.9
LTP/DT	89.5
TPLBP+Chi2	91.6
Gabor+LDA	91.4
IDLS-Distance	93.5
IDLS	97.4
FRID	99.98

从表 1 可以看出:PCA,LDA 和 LARK 的性能并不令人满意,说明这些方法在不同表情和光照变化下的人脸识别并不鲁棒.相比较来说,Gabor 特征、ID-LARK、LBP、TPLBP、LTP 和 IDLS 表现的更鲁棒,且获得了更好的结果.值得注意的是:IDLS 获得了很好的结果,这主要归因于 IDLS 通过图像分解的方法获得了更多的图像局部结构信息.然而在本实验中,IDLS 仍然弱于 FRID,主要原因有 3 点:

- FRID 通过多方向进行图像分解,从而从多角度获得了更丰富的局部结构信息;
- 通过引入一种鲁棒的余弦相关度量方法,进而推导出一种新的图像分解方法,该分解方法实际上综合考虑了方向图像的高频信息和低频信息;
- 通过对高频图像和低频图像分别采用稠密的局部直方图表示,最终得到一种更有判别性的直方图表示方法.

5.2 Extended Yale B 数据库

Extended Yale B 人脸库包含 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本实验中,每张图像的分辨率被调整为 96×84 .本实验为每个人随机选择 K 个图像作为训练集,剩下的图像作为测试集.这里, K 分别取 4,8,12,16,20,24.对于 K 的每种选择,分别随机生成 10 种不同的训练集进行实验,所有方法的平均识别率和标准差被展示在表 2 中.

Table 2 Average recognition rates (percent) and std of different methods on the Extended Yale B database
表 2 在 Extended Yale B 人脸库中不同识别方法的平均识别率和标准差

Methods	4	8	12	16	20	24
PCA	41.6±5.77	59.7±6.26	65.5±7.11	70.6±2.86	75.2±3.09	78.2±4.92
LDA	57.3±11.8	77.5±7.91	86.7±4.10	91.1±2.86	92.3±3.21	94.1±3.38
LARK	84.5±5.30	90.6±2.20	92.2±2.83	92.6±3.20	93.7±3.06	94.9±2.35
ID-LARK	86.8±4.73	91.3±3.43	96.4±1.63	96.9±0.88	97.7±0.54	97.9±0.47
LBP+Chi2	84.2±8.59	93.2±1.68	95.3±1.71	96.3±1.02	96.6±0.91	96.7±1.26
LTP+Chi2	62.1±8.79	76.7±5.08	81.9±3.76	84.5±3.75	85.3±2.52	87.1±3.52
LTP/DT	64.3±10.8	80.8±5.08	86.6±3.73	89.6±3.55	91.1±2.14	92.3±2.61
TPLBP+Chi2	50.3±11.7	66.3±5.67	71.8±5.42	75.4±2.81	76.5±3.36	78.9±5.14
Gabor+LDA	60.3±3.94	78.2±3.98	85.2±3.10	91.2±3.48	93.3±2.83	94.3±1.65
IDLS-Distance	74.9±7.49	84.4±4.16	91.3±2.10	93.5±2.01	93.9±1.69	95.1±1.65
IDLS	95.8±1.73	96.9±0.72	97.9±0.50	98.0±0.66	98.0±0.75	98.2±0.84
FRID	99.52±0.21	99.74±0.04	99.75±0.04	99.76±0.04	99.80±0.05	99.82±0.06

从表 2 可以看出:PCA 和 TPLBP 的识别率相对较低;虽然 LTP/DT,LTP+Chi2,LDA,Gabor+FLDA,LARK 和 IDLS-Distance 的识别率有些提高,但仍不令人满意.ID-LARK,LBP+Chi2 和 IDLS 获得了比以上方法更好的结果.值得注意的是:IDLS 获得了很好的结果,特别当训练样本很少时,IDLS 也获得了很好的结果.IDLS 通过图像分解获得图像更多的局部结构信息,从而获得了更好的结果.然而在本实验中,IDLS 仍然弱于 FRID,FRID 在表 2 的识别结果有两个明显特点:第一,FRID 优于所有的方法,这表明 FRID 在光照变化上表现的更鲁棒;第二,FRID 在不同大小的训练样本下几乎都获得了将近 100% 的识别率,并且每种情况下的标准差都小于 0.3,这表明 FRID 方法表现的更稳定.

5.3 CMU PIE 数据库

CMU PIE 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^[35],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.所有图像的分辨率被设置为 64×64 .

本实验为每个人随机选择 K 个图像作为训练集,剩下的图像作为测试集.这里, K 分别取 5,10,15,20,25.对于 K 的每种选择,分别随机生成 10 种不同的训练集进行实验,所有方法的平均识别率和标准差被展示在表 3 中.

从表 3 可以看出,FRID 获得了最好的识别效果.值得注意的是:当训练样本为 5 时,FRID 获得了远高于其他方法的识别率和更低的标准差;尽管随着训练样本的增多,Gabor,LTP,ID-LARK 和 IDLS 也获得了较好的结果,

但 FRID 仍然获得了近乎完美的结果.和 Extended Yale B 实验中的表现类似,FRID 表现的更稳定.这充分说明了 FRID 特征的鲁棒性.

Table 3 Average recognition rates(percent) and std of different methods on the CMU-PIE database
表 3 在 CMU PIE 人脸库中不同识别方法的平均识别率和标准差

Methods	5	10	15	20	25
PCA	35.9±3.73	53.2±4.83	61.1±7.84	69.2±7.05	78.4±4.67
LDA	59.3±6.56	80.9±4.93	86.7±3.33	89.4±3.26	90.2±2.76
LARK	59.5±6.76	74.6±7.32	82.1±4.63	87.7±5.27	88.5±3.16
ID-LARK	66.9±8.24	89.1±3.67	93.0±2.42	95.4±2.37	96.5±2.00
LBP+Chi2	72.8±4.85	84.0±4.56	90.2±3.38	93.3±3.05	94.8±2.63
LTP+Chi2	81.1±4.64	90.3±3.41	94.0±2.78	96.5±1.82	97.1±1.37
LTP/DT	81.5±4.95	90.4±3.71	94.5±2.43	96.6±1.98	97.4±1.53
TPLBP+Chi2	75.2±6.64	88.8±3.99	93.1±3.12	96.3±2.05	97.4±1.51
Gabor+LDA	75.6±6.93	92.9±2.32	95.4±2.08	96.7±2.00	97.3±1.56
IDLS-Distance	68.1±7.75	87.2±3.24	90.8±3.19	92.7±2.90	93.1±2.33
IDLS	82.6±6.03	94.3±3.89	96.3±2.86	97.4±2.25	97.6±1.91
FRID	96.98±0.73	99.09±0.25	99.77±0.15	99.85±0.10	99.98±0.08

5.4 FRID的参数分析

本节讨论 FRID 方法中不同参数的选择,这里重点讨论多方向计算中的邻居半径 R 、直方图表示中图像分块尺寸 w 和连续的取值区间个数 r .多方向计算中的邻居半径 R 被展示在图 2 中.不同的邻居半径,计算的方向个数不同.当 R 取 1 时,共有 10 个方向被计算(其中 8 个交叉梯度方向,2 个拉普拉斯方向),如图 4 和图 6 所示;当 R 取 2 时,共有 15 个方向被计算(其中 12 个交叉梯度方向,3 个拉普拉斯方向),如图 14 所示(图 14(a)~图 14(i)是交叉梯度方向,其中,图 14(a)表示 4 个交叉梯度方向;图 14(j)~图 14(l)是拉普拉斯方向);当 R 取 3 时,共有 19 个方向被计算(其中 16 个交叉梯度方向,3 个拉普拉斯方向),如图 15 所示(图 15(a)~图 15(m)是交叉梯度方向,其中,图 15(a)表示 4 个交叉梯度方向;图 15(n)~图 15(p)是拉普拉斯方向).

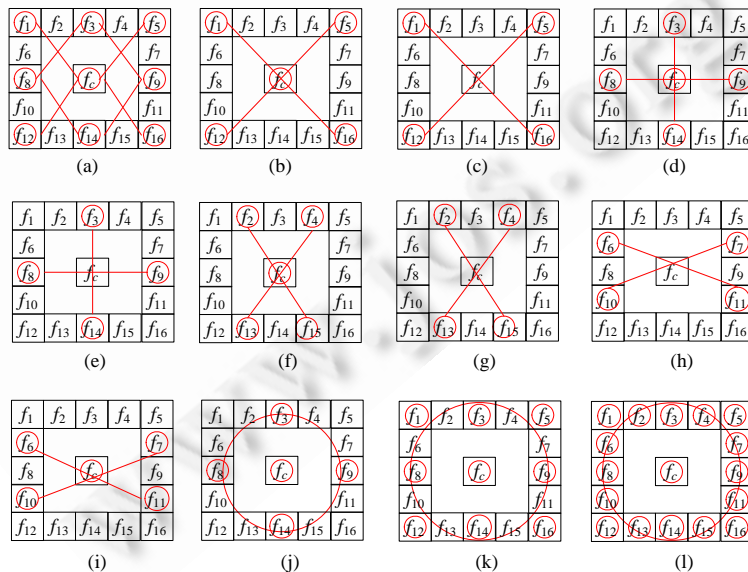


Fig.14 Multiple orientations operator ($R=2$)

图 14 多方向操作($R=2$)

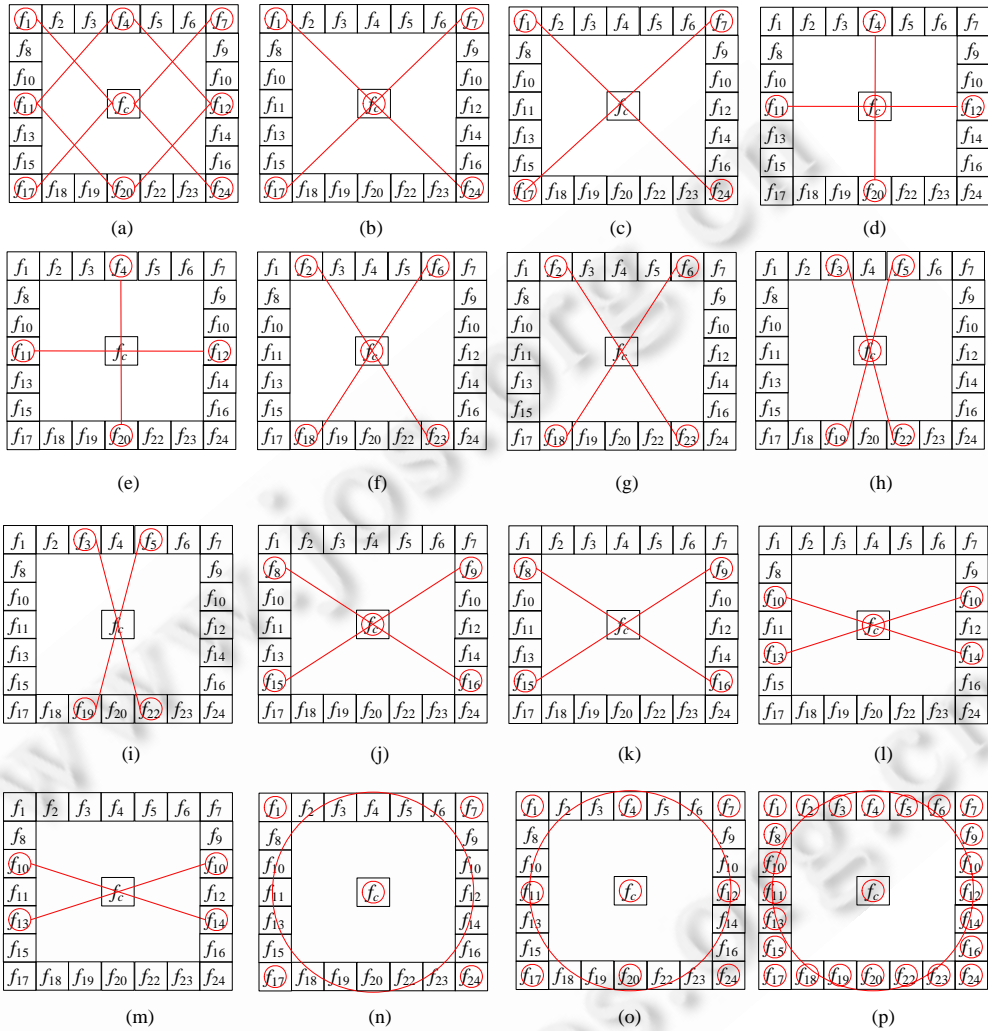


Fig.15 Multiple orientations operator ($R=3$)

图 15 多方向操作($R=3$)

针对不同的邻居半径 R 、不同的分块尺寸 w 和不同的连续取值区间个数 r ,表 4~表 6 分别展示了 3 个数据库上的识别率.从表 4~表 6 可以有如下观察:

- (1) 无论哪一个数据库的实验结果,在相同取值区间 r 和相同邻居半径 R 下,分块尺寸 w 为 3 的识别率都要高于分块尺寸 w 为 5 的识别率.说明分块尺寸 w 取 3 更合适;
- (2) 当区间 r 取 3 时,在不同邻居半径下,不同分块尺寸 w 所对应的识别率分别要高于区间 r 取 2 和 4 所对应的识别结果.说明区间 r 取 3 更合适;
- (3) 当区间 r 取 3、分块尺寸 w 取 3、邻居半径 R 取 2 时,YaleB 和 PIE 数据库都获得最好的识别率.而对于 AR 数据库,当区间 r 取 3、分块尺寸 w 取 3、邻居半径 R 取 1 时,获得最好识别率;但当邻居半径 R 取 2 时,FRID 获得了较好的识别率.

综上所述,针对不同人脸数据库,通常选取的最佳参数配置为:区间 r 取 3,分块尺寸 w 取 3,邻居半径 R 取 2.

Table 4 Recognition rates (percent) of FRID with different parameters on the AR database**表 4** 在 AR 人脸库中,不同参数下 FRID 的识别率

Bins	Block size	Scale size		
		R=1 (%)	R=2 (%)	R=3 (%)
r=2	w=3	99.73	99.67	99.37
	w=5	99.48	99.32	99.20
r=3	w=3	99.98	99.75	99.51
	w=5	99.59	99.60	99.38
r=4	w=3	99.73	99.71	99.26
	w=5	99.55	99.50	99.18

Table 5 Average recognition rates (percent) of FRID with different parameters on the Yale B database**表 5** 在 Yale B 人脸库中,不同参数下 FRID 的平均识别率

Bins	Block size	Scale size		
		R=1 (%)	R=2 (%)	R=3 (%)
r=2	w=3	99.14	99.45	98.04
	w=5	98.71	98.88	97.40
r=3	w=3	99.29	99.52	98.32
	w=5	98.64	98.97	97.94
r=4	w=3	98.52	98.81	97.75
	w=5	98.11	98.16	97.33

Table 6 Average recognition rates (percent) of FRID with different parameters on the PIE database**表 6** 在 PIE 人脸库中,不同参数下 FRID 的平均识别率

Bins	Block size	Scale size		
		R=1 (%)	R=2 (%)	R=3 (%)
r=2	w=3	90.77	92.94	92.16
	w=5	90.17	92.55	91.91
r=3	w=3	95.66	96.98	95.75
	w=5	92.88	93.80	93.28
r=4	w=3	94.27	94.96	94.46
	w=5	92.05	92.87	92.06

6 结束语

本文提出一种基于图像分解的人脸特征表示方法(FRID),该方法首先通过多方向操作把一幅图像分解成一系列方向子图像,这些方向子图像从多个角度刻画了图像的边缘信息和细节信息,从而获得了丰富的局部结构信息;然后,通过非线性欧拉映射把每幅方向子图像分解成实部与虚部图像,分析得知虚部和实部图像分别对应高频和低频信息.欧拉分解实际上综合考虑了方向图像的高频和低频信息.针对实部和虚部图像,分别划分出多个不重叠的局部图像块,通过统计子图像块上不同数值的个数,生成相应的实部和虚部直方图,所有的实部和虚部直方图被串联成一个更有判别性的超级特征向量;最后,通过线性判别分析方法对超级特征向量进行维数约简,以获得每幅输入图像的低维表示.实验结果显示:该方法在多个人脸数据库上获得了优于时新算法的识别结果,并且表现得更加稳定.

致谢 在此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.

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] Turk M, Pentland A. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 1991,3(1):71–86. [doi: 10.1162/jocn.1991.3.1.71]
- [3] 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]
- [4] 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]
- [5] 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]
- [6] Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 2000,290(5500):2323–2326. [doi: 10.1126/science.290.5500.2323]
- [7] 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]
- [8] 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]
- [9] Mika S, Ratsch G, Weston J, Scholkopf B, Smola A, Muller KR. Constructing descriptive and discriminative nonlinear features: Rayleigh coefficients in kernel feature spaces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2003,25(5):623–628. [doi: 10.1109/TPAMI.2003.1195996]
- [10] Yang J, Frangi AF, Yang JY, Zhang D, Jin Z. KPCA plus LDA: A complete kernel fisher discriminant framework for feature extraction and recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2005,27(2):230–244. [doi: 10.1109/TPAMI.2005.33]
- [11] Lowe DG. Distinctive image features from scale invariant key points. *Int'l Journal of Computer Vision*, 2004,60(2):91–110. [doi: 10.1023/B:VISI.0000029664.99615.94]
- [12] Dalal N, Triggs B. Histograms of oriented gradients of human detection. In: *Proc. of the CVPR*. 2005. 886–893. [doi: 10.1109/CVPR.2005.177]
- [13] 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]
- [14] Ojala T, Pietikainen M, Maenpaa 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. [doi: 10.1109/TPAMI.2002.1017623]
- [15] 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. [doi: 10.1109/TIP.2002.999679]
- [16] 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]
- [17] Zhang BC, Gao YS, Zhao SQ, Liu J. Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor. *IEEE Trans. on Image Processing*, 2010,19(2):533–544. [doi: 10.1109/TIP.2009.2035882]
- [18] Zhang WC, Shan SG, Gao W, Chen X, Zhang H. Local gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition. In: *Proc. of the 10th IEEE Int'l Conf. on Computer Vision*. 2005. 786–791. [doi: 10.1109/ICCV.2005.147]
- [19] Zhang BH, Shan SG, Chen XL, Gao W. Histogram of gabor phase patterns (HGPP): A novel object representation approach for face recognition. *IEEE Trans. on Image Processing*, 2007,16(1):57–68. [doi: 10.1109/TIP.2006.884956]
- [20] 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]
- [21] Xie SF, Shan SG, Chen XL, Chen J. Fusing local patterns of gabor magnitude and phase for face recognition. *IEEE Trans. on Image Processing*, 2010,19(5):1349–1361. [doi: 10.1109/TIP.2010.2041397]
- [22] Ahonen T, Rahtu E, Ojansivu V, Heikkila J. Recognition of blurred faces using local phase quantization. In: *Proc. of the Int'l Conf. on Pattern Recognition*. 2008. 1–4. [doi: 10.1109/ICPR.2008.4761847]

- [23] Rahtu E, Heikkilä J, Ojansivu V, Ahonen T. Local phase quantization for blur-insensitive image analysis. *Image and Vision Computing*, 2012,30:501–512. [doi: 10.1016/j.imavis.2012.04.001]
- [24] Lei Z, Ahonen T, Pietikainen M, Li SZ. Local frequency descriptor for low-resolution face recognition. In: *Proc. of the 2011 IEEE Int'l Conf. on Automatic Face & Gesture Recognition and Workshops (FG 2011)*. 2011. 161–166. [doi: 10.1109/FG.2011.5771391]
- [25] Chan CH, Tahir MA, Kittler J. Multiscale local phase quantization for robust component-based face recognition using kernel fusion of multiple descriptors. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2013,35(5):1164–1176. [doi: 10.1109/TPAMI.2012.199]
- [26] Su Y, Shan SG, Chen XL. Hierarchical ensemble of global and local classifiers for face recognition. *IEEE Trans. on Image Processing*, 2009,18(8):1885–1896. [doi: 10.1109/TIP.2009.2021737]
- [27] Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y. Robust face recognition via sparse representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2009,31(2):210–227. [doi: 10.1109/TPAMI.2008.79]
- [28] Zhang L, Yang M, Feng XC. Sparse representation or collaborative representation which helps face recognition? In: *Proc. of the 13th IEEE Int'l Conf. on Computer Vision (ICCV)*. Barcelona, 2011. 471–478. [doi: 10.1109/ICCV.2011.6126277]
- [29] Yang M, Zhang L, Zhang D, Wang S. Relaxed collaborative representation for pattern classification. In: *Proc. of the 25th IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. Providence, 2012. 2224–2231. [doi: 10.1109/CVPR.2012.6247931]
- [30] Qian JJ, Yang J, Xu Y. Local structure-based image decomposition for feature extraction with applications to face recognition. *IEEE Trans. on Image Processing*, 2013,22(9):3591–3603. [doi: 10.1109/TIP.2013.2264676]
- [31] Fitch AJ, Kadyrov A, Christmas WJ, Kittler J. Fast robust correlation. *IEEE Trans. on Image Processing*, 2005,14(8):1063–1073. [doi: 10.1109/TIP.2005.849767]
- [32] Martinez AM, Benavente R. The AR face database. *CVC Technical Report, #24*, 1998.
- [33] 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]
- [34] 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]
- [35] 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]



李照奎(1976—),男,河南邓州人,博士生,副教授,CCF会员,主要研究领域为机器学习,人脸识别,图像分析.
E-mail: lmy52wy@163.com



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.
E-mail: lxding@whu.edu.cn



何进荣(1984—),男,博士生,主要研究领域为机器学习,数据降维.
E-mail: hejinrong@163.com



胡庆辉(1976—),男,博士生,主要研究领域为机器学习,演化计算.
E-mail: huqinghui2004@126.com