

DOI: 10.13382/j.jemi.2017.04.005

局部梯度 DT-CWT 主方向模式的表情识别^{*}

许良凤¹ 王家勇¹ 胡 敏¹ 林 辉² 侯登永¹ 崔婧楠³

(1. 合肥工业大学 情感计算与先进智能机器安徽省重点实验室 合肥 230009;
2. 合肥工业大学 电子科学与应用物理学院 合肥 230009; 3. 华南理工大学 广州 510006)

摘 要:提出一种采用局部梯度双树复小波变换(dual-tree complex wavelet transform, DT-CWT)主方向模式(dominant direction pattern, DDP)的人脸表情识别方法。首先,对归一化后的表情图像进行4层DT-CWT,每一层得到8个方向的DT-CWT特征图像,其中包括6个高频方向和2个低频方向,构建一种新的主方向模式(IDDP)对每个DT-CWT特征图像进行编码;然后按照基于梯度方向的融合规则将每一层上IDDP编码特征图融合到一起,将融合图进一步划分为若干个不重叠且大小相等的子块,分别计算每个子块区域的直方图分布,将其联合起来得到人脸表情图像的特征;最后,采用基于Fisher加权的Chi平方概率统计最近邻方法进行分类识别。大量实验表明,算法在识别率和识别时间上都体现出了一定的优势。

关键词:表情识别;双数复小波变换;主方向模式;梯度方向;分块直方图

中图分类号:TP391;TN911 **文献标识码:**A **国家标准学科分类代码:**510.6040

Facial expression recognition based on local gradient DT-CWT dominant direction pattern

Xu Liangfeng¹ Wang Jiayong¹ Hu Min¹ Lin Hui² Hou Dengyong¹ Cui Jingnan¹

(1. Anhui Province Key Laboratory of Affective Computing and Advanced Intelligent Machine, Hefei University of Technology, Hefei 230009, China; 2. School of Electronic Science and Applied Physics, Hefei University of Technology, Hefei 230009, China; 3. South China University of Technology, Guangzhou 510006, China)

Abstract: A novel facial expression recognition is proposed in the paper, in which the local gradient dual-tree complex wavelet transform dominant direction pattern is used. Firstly, four layers DT-CWT are used on normalized expression image. For each layer, we can obtain the feature images of eight directions, which include 6 high-frequency directions and 2 low-frequency directions. A new DDP (IDDP) is constructed, and which is used to code for each DT-CWT feature image. Secondly, the IDDP feature images of each layer in different directions are fused based on rules of gradient direction, and every fused image is divided into several non-overlapping and equal-sized blocks. The corresponding histogram of the fused feature in each block is calculated respectively, and the final feature of facial expression image is obtained by cascading all of them. Finally, the nearest neighbor method based on Chi Square statistic weighted by Fisher is used to classify and identify. A large number of experiments show that the proposed method has a certain advantage on the recognition rate and recognition time.

Keywords: expression recognition; dual-tree complex wavelet transform; dominant direction pattern; gradient direction; block histogram

1 引 言

人脸表情可以体现人们的情感变化,在人们交流过程

中,它起到一定的信息传递作用。从心理学角度上来说,甚至可以通过人脸表情推测出人们的内心活动。近年来,随着人工智能、图像分析和模式识别等领域的发展,人脸表情识别已经成为模式识别领域的热点之一^[1-4]。

人脸表情识别系统主要由3个模块组成:预处理模块、特征提取模块和分类识别模块^[5],其中,特征提取和分类识别是至关重要的。衡量人脸表情识别系统性能最主要的两个因素是识别率和时间复杂度。高识别率要求在特征提取模块能完整的提取出表情图像的本质特征,在分类识别模块能设计出性能优越的分类器;低时间复杂度的实现在于较低的特征维数和原理简单的分类器。文献[6]提出一种径向局部 Gabor 特征编码和分类器合成的人脸表情识别方法。该方法采用径向网络对 Gabor 多尺度局部特征进行联合编码,虽然降低了特征维数,但在编码中部分原始细节信息丢失将导致局部特征描述不充分,降低识别系统的识别率。文献[7]提出一种融合全局与局部多样性特征的人脸表情识别方法(GLDPE),采用 PCA 和流形学习提取全局结构特征和局部流形结构多样性特征,将全局特征与局部多样性特征相结合,提取出低维流形特征用于表情识别。该方法不仅充分保持了样本的全局特性,而且有效地提取出局部流形结构的差异性和相似性,但实验中参数 K 是通过实验得出的结果,是相对耗时的。文献[8]提出一种多特征多分类器优化匹配的人脸表情识别方法,首先构造多特征-多分类器最佳匹配的自适应人脸表情识别模型,选定特征和分类器,模型根据分类器的准确性和结果的多样性选择最优的分类器组合,针对多分类器投票系统不能识别的表情图像,通过自适应分类决策来识别。虽然本文获得了较高的识别率,但是时间复杂度较高。文献[9]提出一种基于纹理和形状的人脸表情识别方法,利用 AAM 提取表情图像的形状和纹理特征,然后通过判断保留的信息来降低特征的维数,该方法的鲁棒性和实时性较好,但识别率一般。

综上所述表明,在高识别率的人脸表情识别系统中,时间复杂度常常较高,在时间复杂度较低的人脸表情识别系统中,识别率常常较差。针对此问题,本文在文献[10-13]的启发下:1)利用 DT-CWT 提取图像的特征,DT-CWT 具有近似的平移不变性和良好的方向选择性,可以从多方向多层次上提取图像的纹理信息^[14],本文对 DT-CWT 进行了改进,不仅提取了图像的6个高频方向特征,同时还保留了2个低频方向特征,把高频和低频特征结合起来形成更加全面的图像特征;2)构建一种新的主方向模式(improve dominant direction pattern, IDDP)对 DT-CWT 特征进行编码,与 DDP 相比, IDDP 结合了更多邻域的方向信息,编码后的特征具有更高的区分性;3)采用基于梯度方向的融合规则将每一层上不同方向的 IDDP 编码图融合到一起,提取融合图的分块直方图序列。融合特征^[15-16]既能有效的降低特征数据间的冗余,又能保留图像基本的有效信息,提高了系统的实时性;另分块直方图既提取了融合图像的全

局特征,又保护了融合图像的结构细节,提高了系统的准确性。

2 局部 DT-CWT 主方向模式的表情特征提取方法

2.1 DT-CWT 特征

为了克服传统小波变换的缺陷,Kingsbury^[17]提出了一种新的小波变换:双树复小波变换。该方法与常规的小波变换不同,由2个平行的小波变换树组成,2个小波变换树都由传统的实数滤波器构成,2个小波树分别产生复小波变换的实部和虚部^[18]。一维信号的 DT-CWT 表示如下:

$$\psi(t) = \psi_h(t) + j\psi_g(t) \quad (1)$$

式中: $\psi(t)$ 为复小波, $\psi_h(t)$ 为树 a 的小波, $\psi_g(t)$ 为树 b 的小波。

对于二维图像信号,需要用二维的 DT-CWT,即分别对图像的行和列进行一维 DT-CWT。相对于传统小波,DT-CWT 具有更好的方向选择性,在对图像进行每一层分解时,DT-CWT 都会在高频信号的6个方向($\pm 15^\circ$ 、 $\pm 45^\circ$ 、 $\pm 75^\circ$)上采样,获取6个高频子图像,因此能够获得很丰富的纹理信息^[19]。图1展示了对人脸表情图像进行4层 DT-CWT 后得到的高频子图像。如图1所示,3种特征中模特征的纹理信息比较丰富,因为实部特征和虚部特征仅仅包含了图像的一部分信息,而模特征结合了实部和虚部,包含了更全面的人脸表情图像信息。故本文选用模特征来表征人脸图像信息。

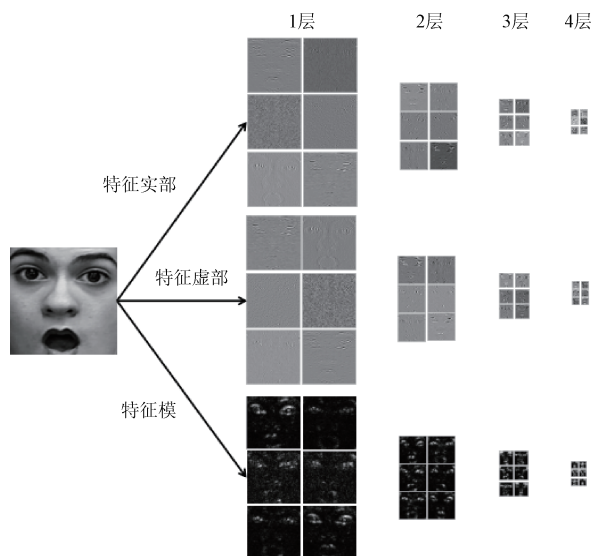


图1 人脸表情图像的4层 DT-CWT 高频特征

Fig. 1 4-layer DT-CWT high frequency features of a facial expression images

2.2 DDP 和 IDDP

DDP 编码后的图像不仅具有丰富的纹理信息,而且还在一定程度上降低了特征维数^[13],DDP 算子的邻域大小为 3×3,每一个像素点的周围 8 邻域分别对应着由 3 比特二进制数构成的 8 个方向数,如图 2 所示。DDP 编码的基本思想为比较中心像素点的 8 邻域像素值,选出最大值和次大值作为主要像素进行编码,如果 2 个或 2 个以上的邻域像素拥有相同的最大值或次大值,选择较小方向数上的像素作为最大值或次大值。DDP 编码规则为 $\min(D_1 \times 8 + D_2, D_2 \times 8 + D_1)$,其中 D_1 和 D_2 分别是最大像素值和次大像素值所对应的方向数,DDP 编码示例如图 3(a)所示。DDP 编码后灰度值范围为 1~55,相对于 LBP、LGC 和 LDC 等一些经典算子,DDP 算子在提取丰富纹理信息的基础上有效的降低了特征数据间的冗余。

由于 DDP 在处理拥有相同最大值和次大值的邻域像素时,只选择较小方向数上的像素,而舍弃其他方向数上的像素,造成图像特征的丢失。故本文对 DDP 进行了改进,把那些拥有相同最大值和次大值的邻域像素都作为主要像素进行编码。编码规则为:

$$\min\left\{\left\lceil 8 \times \sum_{i=1}^N D_i / N \right\rceil + \left\lceil \sum_{j=N+1}^{N+M} D_j / M \right\rceil, \left\lceil 8 \times \sum_{j=1}^M D_j / M \right\rceil + \left\lceil \sum_{i=M+1}^{N+M} D_i / N \right\rceil\right\}。$$

其中: D_i 、 D_j 分别表示邻域像素最大值和次大值所对应的方向数, N 、 M 分别表示拥有相同最大值和次大值的邻域像素个数, $N \in [1,8]$, $M \in [0,7]$, $M + N \in [1,8]$, $\lceil \cdot \rceil$ 表示向上取整。IDDP 编码示例如图 3(b)所示。

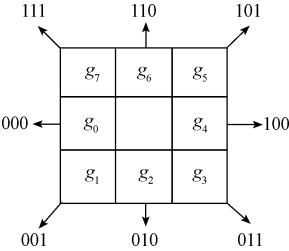


图 2 DDP 的 3×3 大小邻域的示意图
Fig. 2 Schematic diagram of 3×3 template of DDP operator

与 DDP 相比, IDDP 结合了更多邻域的方向信息,从而使得编码后的特征具有更高的区分性。如图 4 所示,用 DDP 和 IDDP 分别对 4 幅灰度分布不同的 3×3 图像进行编码,4 幅图像的 DDP 编码值相同,而 IDDP 编码值却各不相同,所以 IDDP 特征具有更高的鉴别能力。人脸表情图像 DT-CWT 特征的 IDDP 编码图如图 5 所示。

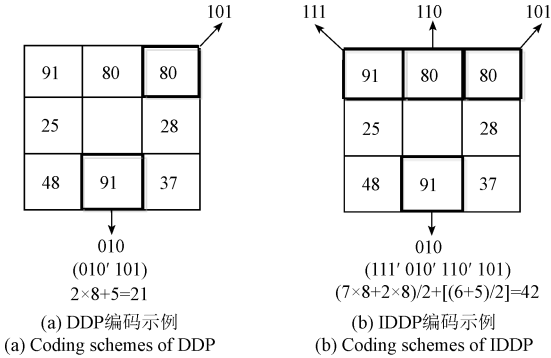


图 3 DDP 和 IDDP 编码示例
Fig. 3 Coding examples of DDP and IDDP

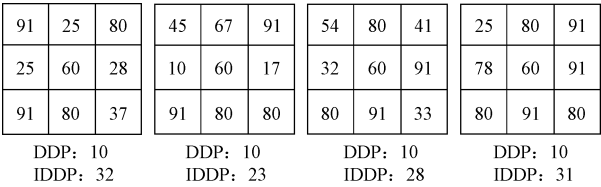


图 4 不同灰度分布示意图
Fig. 4 Schematic diagram of different gray distribution

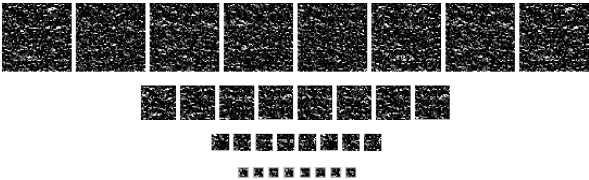


图 5 4 层 8 方向 DT-CWT 特征的 IDDP 编码图
Fig. 5 IDDP coding map of DT-CWT features on 4 layers and 8 directions

2.3 基于梯度方向的融合规则

本文对表情图像进行 4 层 DT-CWT,每一层得到 8 个子图像。对于人脸表情而言,虽然提取图像多方向多层级的纹理信息非常重要,但是 DT-CWT 会使每张表情图像转化成 32 个对应不同层级和方向的图像,经过 IDDP 编码后特征维数仍然远远高于原始图像,造成特征数据冗余。为了降低特征数据的冗余性,本文构建了一种基于梯度方向的融合规则,将同一层不同方向的 IDDP 特征按照梯度方向进行融合。

按照图 6 所示的组合方式,将 IDDP 特征 $\{X_{u,v}(z) | v \in (1, \cdots, 8), u \in (1, \cdots, 4)\}$ 与融合之后的特征 $\{Y_u(z) | u \in (1, \cdots, 4)\}$ 进行排列组合。其中 u 表示层数, v 表示方向, $\{X_{u,v}(z)\}$ 表示第 u 层第 v 个方向上 IDDP 特征, $\{Y_u(z)\}$ 表第 u 层融合后的特征。

需要选择合适的梯度算子来计算图 6 中组合特征的梯度方向。梯度算子分为两种:一阶微分梯度算子和二

$X_{u,1}(z)$	$X_{u,8}(z)$	$X_{u,4}(z)$
$X_{u,2}(z)$	$Y_u(z)$	$X_{u,5}(z)$
$X_{u,3}(z)$	$X_{u,7}(z)$	$X_{u,6}(z)$

图6 IDDP 特征的组合方式

Fig.6 The combinations mode of IDDP features

阶微分梯度算子。从数学的层面上来说,一阶微分反映的是变化率,二阶微分反映的是变化率的变化率。本文需要计算 IDDP 特征的梯度方向,梯度方向是数据变化最快的方向,即变化率最快,而一阶微分恰好体现了变化率。常用的一阶微分梯度算子有 Robers 算子、Sobel 算子、Prewitt 算子。其中,Robers 算子模板大小为 2×2 ,与本文 8 个方向 IDDP 特征无法匹配;Sobel 算子增加了部分方向特征的权重,导致特征融合产生误差;Prewitt 算子不仅模板大小与本文匹配,而且 8 个方向特征的权重均为 1,特征融合时不会造成梯度方向的偏差。故本文选 Prewitt 算子。通过 Prewitt 算子计算可得:

$$G_x = (X_{u,3}(z) + X_{u,7}(z) + X_{u,6}(z)) - (X_{u,1}(z) + X_{u,8}(z) + X_{u,4}(z)) \quad (2)$$

$$G_y = (X_{u,4}(z) + X_{u,5}(z) + X_{u,6}(z)) - (X_{u,1}(z) + X_{u,2}(z) + X_{u,3}(z)) \quad (3)$$

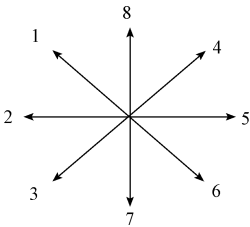
其中 G_x 表示水平方向的偏导数, G_y 表示垂直方向的偏导数。由式(2)和(3)可得梯度算子的方向为:

$$\theta = \arctan(G_y/G_x) \quad (4)$$

如图 7 所示,将 θ 所对应的梯度方向作为 $Y_u(z)$ 的值,则:

$$Y_u(z) = \begin{cases} N + 4, & N \leq 4 \\ N - 4, & N \geq 5 \end{cases} \quad (5)$$

式中: $N = \lceil \theta/45^\circ \rceil$, $\lceil \cdot \rceil$ 表示返回不大于参数的最大整数值。

图7 θ 与梯度方向的对应关系Fig.7 The corresponding relationship between the θ and gradient direction

经过融合后,一副表情图像形成 4 个层级上的 IDDP 融合图,如图 8 所示。



图8 4 个层级上的 IDDP 融合图

Fig.8 IDDP fusion images on 4 layers

2.4 分块直方图

由于直方图在提取特征时只考虑了图像的全局分布,而忽略了局部分布,导致直方图特征缺乏局部表征能力。分块直方图首先对图像进行分块,然后再提取直方图特征,最后将各个子块直方图序列串联。分块直方图特征弥补了直方图特征的缺点,不仅具有很好的全局表征能力,而且还包含了一些局部区域内的信息,能够兼顾局部和整体^[20]。因此,本文提取 IDDP 融合图的分块直方图序列。

分块直方图的工作过程如下:

1) 将每一层的 IDDP 融合图划分为 m 个均匀且不重叠的子区域,分别表示为 $R_0, \dots, R_{m-1}$ 。由于每层融合图的大小不同,所以 m 的值随着层数的变化而变化;

2) 分别统计每一层子区域 $R_0, \dots, R_{m-1}$ 的直方图分布,并将同一层上子区域的直方图依次串联;

3) 将 4 个层级上分块直方图序列串接到一起,得到 IDDP 融合图的分块直方图序列。

3 局部梯度 DT-CWT 主方向模式的表情识别方法

Chi 平方概率统计是一种简单而有效的直方图相似度度量方法,其计算公式为:

$$\chi^2(H^1, H^2) = \sum_{i=1}^l \frac{(H_i^1 - H_i^2)^2}{H_i^1 + H_i^2} \quad (6)$$

式中: H^1 和 H^2 表示直方图序列, l 表示直方图的维度。式(6)通过计算直方图 H^1 和 H^2 之间的距离来衡量两直方图之间相似性的高低。利用这种度量方式来计算样本间分块直方图的相似性,计算公式为:

$$S(R^1, R^2) = \sum_{u=1}^4 \sum_{r=0}^{m-1} \chi^2(H_{u,r}^1, H_{u,r}^2) \quad (7)$$

由于人脸图像的不同区域对表情识别的贡献率存在差异,如眉毛、眼睛和嘴巴等区域具有较强的纹理信息,而鼻子、面部等区域所包含的纹理信息较少。所以,为了增强表情特征的区分性,运用 Fisher 准则对人脸图像的不同区域赋予相应的权值,加权后的相似性度量公式为:

$$S'(R^1, R^2) = \sum_{u=1}^4 \sum_{r=0}^{m-1} \omega_{u,r} \chi^2(H_{u,r}^1, H_{u,r}^2) \quad (8)$$

式中: $\omega_{u,r}$ 是第 u 层 IDDP 融合图中第 r 块区域的权值。Fisher 准则是通过使类内散度尽可能小,而类间散度尽

可能大来分配每个分块的权值。对于 C 分类问题,同类样本中不同样本之间的相似度分布构成了类内相似度空间,不同类样本中样本之间的相似度分布构成了类间相似度空间,则类内相似度均值为:

$$m_{u,r}^I = \frac{1}{C} \sum_{i=1}^C \frac{2}{N_i(N_i-1)} \sum_{k=2}^{N_i} \sum_{j=1}^{k-1} \chi^2(H_{u,r}^{(i,j)}, H_{u,r}^{(i,k)}) \tag{9}$$

式中: N_i 是第 i 类样本的个数, $H_{u,r}^{(i,j)}$ 表示第 i 类中第 j 个样本的第 u 层 IDDP 融合图中第 r 块区域的直方图。类内相似度方差为:

$$(\sigma_{u,r}^I)^2 = \frac{2}{\sum_{i=1}^C N_i(N_i-1)-2} \times \sum_{i=1}^C \sum_{k=2}^{N_i} \sum_{j=1}^{k-1} [\chi^2(H_{u,r}^{(i,j)}, H_{u,r}^{(i,k)}) - m_{u,r}^I]^2 \tag{10}$$

同理,类间相似度均值为:

$$m_{u,r}^E = \frac{2}{C(C-1)} \sum_{i=1}^{C-1} \sum_{j=i+1}^C \frac{1}{N_i N_j} \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} \chi^2(H_{u,r}^{(i,k)}, H_{u,r}^{(j,l)}) \tag{11}$$

类间相似度方差为:

$$(\sigma_{u,r}^E)^2 = \frac{1}{\sum_{i=2}^C \sum_{j=1}^{i-1} N_i N_j - 1} \times \sum_{i=1}^{C-1} \sum_{j=i+1}^C \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} [\chi^2(H_{u,r}^{(i,k)}, H_{u,r}^{(j,l)}) - m_{u,r}^E]^2 \tag{12}$$

最终可得 $\omega_{u,r}$ 为:

$$\omega_{u,r} = \frac{(m_{u,r}^I - m_{u,r}^E)^2}{(\sigma_{u,r}^I)^2 + (\sigma_{u,r}^E)^2} \tag{13}$$

本文利用 Fisher 加权的 Chi 平方概率统计来计算待测样本与训练样本之间的距离,并通过最近邻分类器进行识别。

4 实验结果及分析

4.1 实验描述

为了测试本文算法的性能,实验选用比较常用的表情库:日本女性人脸表情库(JAFFE)和 Cohn-Kanade(CK)库,在 Windows XP 系统上(单核 2.8 GHz CPU, 2.0 GB 内存)用 MATLAB 实现测试。

本文算法的步骤如下:

- 1) 预处理,灰度归一化和尺度归一化,预处理后图片大小为 128×128 像素,如图 9 所示;
- 2) 提取归一化后人脸图像的 4 层 DT-CWT 特征,并对其进行 IDDP 编码;
- 3) 按照基于梯度方向融合规则将每层上的 IDDP 特征进行融合,并提取融合图的分块直方图序列;
- 4) 采用 Fisher 加权的 Chi 平方概率统计最近邻方法进行分类识别。

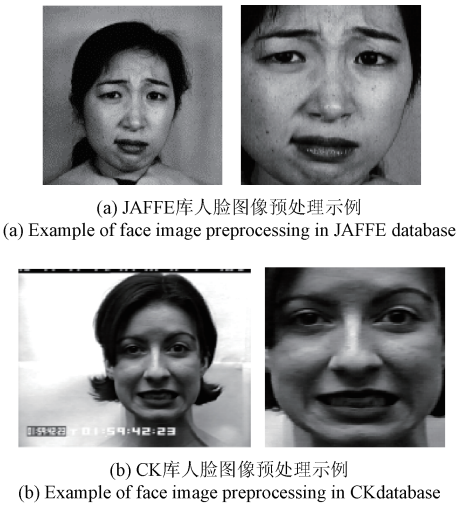


图 9 预处理示例
Fig. 9 Preprocessing image samples

4.2 实验相关参数选取及实验结果分析

本文所有的实验都是在 JAFFE 库和 CK 库上完成的。在 JAFFE 中,随机选取每个人的每种表情图像 1~2 幅作为训练样本,剩下的作为测试样本,每次实验训练样本数为 139,测试样本数为 74;在 CK 库中,首先选取 10 个人的表情作为样本,然后随机选取每个人的每种表情 1~3 幅图像作为训练样本,在每个人的剩余图像中随机选取每种表情 1~3 幅图像作为测试样本,每次实验训练样本数为 105,测试样本数为 105。每个库上进行 3 组交叉实验,计算出平均识别时间和识别率。

4.2.1 DT-CWT 层数的选取

DT-CWT 是对人脸表情图像进行逐层变换得到表情特征。若层数较少,图像的纹理信息无法较完全地从低频图像中提取出来,造成特征缺失;若层数较多,虽然纹理信息较完整,但特征的维数会很高,造成特征冗余,增加计算复杂度。为了保证系统的准确性和实时性,必须选择较合适的变换层数。在 JAFFE 库和 CK 库上对 DT-CWT 层数的选取进行研究,实验结果如表 1 所示。

表 1 JAFFE 库和 CK 库上不同 DT-CWT 层数的识别时间和识别率比较

Table 1 The comparison of recognition rate and recognition time of different DT-CWT layer on JAFFE and CK database

DT-CWT 层数	JAFFE 库		CK 库	
	识别时间/ms	识别率/%	识别时间/ms	识别率/%
1	1 750	81.08	2435	88.57
2	1 940	89.19	2775	93.33
3	2 030	94.59	2885	97.14
4	2 050	96.85	2910	98.41
5	2 055	96.85	2920	98.41

从表1中可以看出,JAFFE库和CK库的实验结果大致相同,随着DT-CWT层数的增加,识别时间和识别率都在逐渐提高,但提高的越来越缓慢。首先,从识别率上来看,在1~3层之间每增加一层,识别率有明显的提高,从第3~4层识别率的提高变得缓慢,从第4~5层识别率没有再提高,说明随着DT-CWT层数的增加,图像纹理信息逐渐饱和,到第4层时达到了饱和状态;从识别时间上来看,随着DT-CWT层数的增加,识别时间一直在增加,从第4~5层,在识别率没有提高的情况下,识别时间仍在缓慢提高。综上,本文选用4层DT-CWT来提取表情图像的特征。

4.2.2 分块直方图分块大小的选取

分块直方图不仅具有很好的全局表征能力,而且还包含了一些局部区域内的信息,能够兼顾局部和整体。所以分块大小的选取至关重要,若分块太小,图像的结构细节虽然得到了保护,但图像的全局特征却遭到了破坏;相反,若分块太大,全局特征得到了保护,结构细节却遭到破坏。本文在JAFFE库和CK库上对直方图分块大小进行了一系列的实验,实验结果如表2、3所示。

表2 JAFFE库上不同分块大小的识别率比较
Table 2 The comparison of recognition rate of different block size on JAFFE database

DT-CWT 层数	分块大小			
	2×2	4×4	8×8	16×16
1	58.10	72.97	81.08	62.16
2	68.92	81.08	89.19	74.32
3	77.03	87.84	94.59	82.43
4	81.34	89.19	96.85	-

表3 CK库上不同分块大小的识别率比较
Table 3 The comparison of recognition rate of different block size on CK database

DT-CWT 层数	分块大小			
	2×2	4×4	8×8	16×16
1	75.36	84.95	88.57	73.25
2	79.57	88.52	93.33	77.18
3	83.03	90.28	97.14	80.75
4	85.41	92.35	98.41	-

从表2、3中可以看出,在两个库上无论进行几层DT-CWT都有一个相同的规律:当分块大小从2×2增大到8×8时,识别率逐渐提高,但若继续增大到16×16时,识别率又开始下降。两个表中都是分块大小为8×8时识别率达到最大值。故本文分块直方图的分块大小为8×8。

4.2.3 本文算法实验结果分析

表4和5给出了本文算法在JAFFE库和CK上的实

验结果,可以看出在JAFFE库和CK库上平均识别率分别为96.85%和98.41%。其中,JAFFE库上有5种表情出现了误识别,CK库上有4种表情出现了误识别,从表情图片的特征中可以看出,误识别表情之间的纹理信息相似度很高,误识别示例如图10所示。

表4 基于JAFFE人脸表情库的实验结果
Table 4 The experiment results based on JAFFE database

表情	测试 样本数	正确识别数			平均 识别率/%
		第1次	第2次	第3次	
中性	10	10	10	10	100
高兴	11	11	11	11	100
生气	10	10	9	10	96.67
悲伤	11	10	10	11	93.94
恐惧	11	10	11	10	93.94
惊讶	11	11	11	10	96.97
厌恶	10	10	10	9	96.67
总计	74	72	72	71	96.85

表5 基于CK人脸表情库的实验结果
Table 5 The experiment results based on CK database

表情	测试 样本数	正确识别数			平均 识别率/%
		第1次	第2次	第3次	
中性	15	15	15	15	100
高兴	15	15	15	15	100
生气	15	15	14	15	97.78
悲伤	15	14	15	14	95.5
恐惧	15	14	15	15	97.78
惊讶	15	15	15	15	100
厌恶	15	15	15	14	97.78
总计	105	103	104	103	98.41

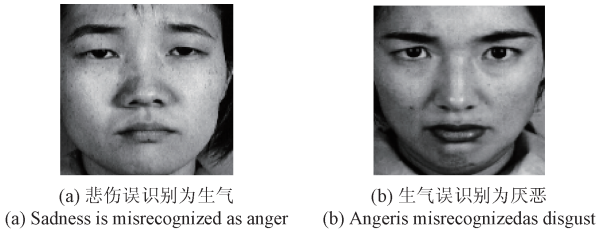


图10 误识别表情示例
Fig. 10 Mistaken images samples

4.2.4 不同算法的识别率和识别时间对比

本文在JAFFE库和CK库上比较了几种人脸表情识别算法与本文算法的性能,实验结果如表6、7所示。

从表6、7中可以看出:1)本文算法不仅识别率高,而且在识别时间上也有一定的优势,具有较高的准确性和较低的时间复杂度;2)采用DDP编码得到的识别率低于LBP编码,而IDDP编码后的识别率却高于LBP编码,表明了IDDP编码的有效性;3)无论采用DDP、LBP还

表 6 JAFFE 库上不同算法识别率和识别时间比较
Table 6 The comparison of recognition rate and recognition time of different algorithm on JAFFE database

算法	算法性能	
	识别率/%	识别时间/ms
Gabor + Radial Encoding ^[6]	89.67	2 310
GLDPE ^[7]	83.67	3 379
多特征多分类器 ^[8]	94.77	28 140
Shape + Texture ^[9]	93.25	3 580
Gabor + LBP	97.1	33 830
DT-CWT + DDP + 融合	87.84	1 950
DT-CWT + LBP + 融合	90.05	2 250
DT-CWT + IDDP + 融合	90.54	1 960
DT-CWT + DDP + 融合(W)	93.24	2 040
DT-CWT + LBP + 融合(W)	95.94	2 380
DT-CWT + IDDP + 融合(W)	96.85	2 050

表 7 CK 库上不同算法识别率和识别时间比较
Table 7 The comparison of recognition rate and recognition time of different algorithm on CK database

算法	算法性能	
	识别率/%	识别时间/ms
Gabor + Radial Encoding ^[6]	91.51	3 280
GLDPE ^[7]	97.08	3 128
多特征多分类器 ^[8]	97.21	38 643
Shape + Texture ^[9]	95.94	5 080
Gabor + LBP	97.81	48 000
DT-CWT + DDP + 融合	90.48	2 770
DT-CWT + LBP + 融合	92.86	3 190
DT-CWT + IDDP + 融合	94.29	2 780
DT-CWT + DDP + 融合(W)	95.24	2 890
DT-CWT + LBP + 融合(W)	97.14	3 380
DT-CWT + IDDP + 融合(W)	98.41	2 910

是 IDDP 编码, Fisher 加权后识别率都明显提高, 表明加权后的表情特征区分性更高。

5 结 论

本文提出了一种局部梯度 DT-CWT 主方向模式的人脸表情识别方法。大量实验表明, 该方法不仅在识别率上有所提高, 而且在识别时间上也体现出了如下优越性。

1) DT-CWT 具有近似的平移不变性和良好的方向选择性, 可以从多方向多层次上提取图像的纹理信息, 能够提取比较完整的纹理特征。

2) 对 DT-CWT 进行了改进, 不仅提取了图像的高频特征, 同时还保留了低频特征, 把高频特征和低频特征结合起来得到了更加全面的图像特征。

3) 提出了 IDDP 算子, 相对于 DDP, IDDP 结合了更多邻域信息, 具有更高的区分性和鉴别能力。

4) 构建了基于梯度方向的融合规则, 融合特征既能有效地降低特征数据间的冗余, 又能保证决策信息不会丢失, 降低了系统的时间复杂度。

5) 提取 IDDP 融合图的分块直方图, 分块直方图既提取了融合图像的全局特征, 同时又保护了融合图像的结构细节, 提高了系统的准确性。

参考文献

[1] 胡步发, 王金伟. 双模态及语义知识的三维人脸表情识别方法[J]. 仪器仪表学报, 2013, 34(4):873-880.
HU B F, WANG J W. 3D facial expression recognition method based on bimodal and semantic knowledge[J]. Chinese Journal of Scientific Instrument, 2013, 34(4): 873-880.

[2] ELEFThERIADIS S, RUDOVIC O, PANTIC M. Discriminative shared Gaussian processes for multiview and view-invariant facial expression recognition[J]. IEEE Transactions on Image Processing, 2015, 24(1): 189-204.

[3] 於俊, 汪增福, 李睿. 一种同步人脸运动跟踪与表情识别算法[J]. 电子学报, 2015, 43(2):371-376.
YU J, WANG Z F, LI R. A simultaneous facial motion tracking and expression recognition algorithm[J]. Acta Electronica Sinica, 2015, 43(2): 371-376.

[4] ZHANG W, ZHANG Y, MA L, et al. Multimodal learning for facial expression recognition[J]. Pattern Recognition, 2015, 48(10):3191-3202.

[5] 胡敏, 李堃, 王晓华, 等. 基于直方图加权 HCBP 的人脸表情识别[J]. 电子测量与仪器学报, 2015, 29(7):953-960.
HU M, LI K, WANG X H, et al. Facial expression recognition based on histogram weighted HCBP[J]. Journal of Electronic Measurement and Instrumentation, 2015, 29(7):953-960.

[6] GU W, XIANG C, VENKATESH Y V, et al. Facial expression recognition using radial encoding of local Gabor features and classifier synthesis[J]. Pattern Recognition, 2012, 45(1):80-91.

[7] 李雅倩, 李颖杰, 李海滨, 等. 融合全局与局部多样性特征的人脸表情识别[J]. 光学学报, 2014, 34(5): 1-7.
LI Y Q, LI Y J, LI H B, et al. Fusion of global and local various feature for facial expression recognition[J]. Acta Optica Sinica, 2014, 34(5):1-7.

[8] 王晓华, 黄伟, 金超, 等. 多特征多分类器优化匹配的人脸表情识别[J]. 光电工程, 2016, 43(3):73-79.
WANG X H, HUANG W, JIN CH, et al. Facial expression recognition based on the optimal matching of

- multi-feature and multi-classifier [J]. Opto-Electronic Engineering, 2016, 43(3):73-79.
- [9] ZHENG W, LIU C. Facial expression recognition based on texture and shape [C]. 25th Wireless and Optical Communication Conference (WOCC), IEEE, 2016: 1-5.
- [10] 刘帅师, 田彦涛, 万川. 基于 Gabor 多方向特征融合与分块直方图的人脸表情识别方法 [J]. 自动化学报, 2011, 37(12):1455-1463.
- LIU SH SH, TIAN Y T, WAN C. Facial expression recognition method based on Gabor multi-orientation features fusion and block histogram [J]. Acta Automatica Sinica, 2011, 37(12):1455-1463.
- [11] 马本学, 高国刚, 王宝, 等. 基于双树复小波变换和邻域操作的哈密瓜纹理提取 [J]. 农业机械学报, 2014, 45(12):316-322.
- MA B X, GAO G G, WANG B, et al. Texture extraction of Hami Melon on dual tree complex wavelet transform and neighborhood operation [J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(12): 316-322.
- [12] 王仕民, 叶继华, 邓涛, 等. 2 维双树复小波不确定度加权融合的人脸识别 [J]. 中国图象图形学报, 2012, 17(8):995-1001.
- WANG SH M, YE J H, DENG T, et al. Two-dimensional dual-tree complex wavelet transform uncertainty weighted fusion in face recognition [J]. Journal of Image & Graphics, 2012, 17(8):995-1001.
- [13] LIU J, JING X, SUN S, et al. Local Gabor dominant direction pattern for face recognition [J]. Chinese Journal of Electronics, 2015, 24(2):245-250.
- [14] 宋怡焕, 饶秀勤, 应义斌. 基于 DT-CWT 和 LS-SVM 的苹果果梗/花萼和缺陷识别 [J]. 农业工程学报, 2012, 28(9):114-118.
- SONG Y H, RAO X Q, YING Y B. Apple stem/calyx and defect discrimination using DT-CWT and LS-SVM [J]. Transactions of the Chinese Society of Agricultural Engineering, 2012, 28(9): 114-118.
- [15] 杨飞, 苏剑波. 人脸显性特征的融合构造方法及识别 [J]. 电子学报, 2012, 40(3):466-471.
- YANG F, SU J B. Face recognition based on explicit facial features by fusion construction method [J]. Acta Electronica Sinica, 2012, 40(3):466-471.
- [16] 孙东辉, 鞠秀亮, 冯登超, 等. 基于 FAST 检测器和 SURF 描述子的聚合图像人脸识别 [J]. 国外电子测量技术, 2016, 35(1):94-98.
- SUN D H, JU X L, FENG D CH, et al. Aggregated image face based on Fast detection and SURF descriptor [J]. Foreign Electronic Measurement Technology, 2016, 35(1):94-98.
- [17] KINGSBURY N. Image processing with complex wavelets [J]. Philosophical Transactions of the Royal Society of London A Mathematical Physical & Engineering Sciences, 1999, 357(1760):2543-2560.
- [18] 蔡蕾, 张春雨, 李斌, 等. 基于 Q-shift 双树复数小波系数的纹理图像检索 [J]. 光电子·激光, 2009, 20(9):1252-1257.
- CAI L, ZHANG CH Y, LI B, et al. Texture image retrieval based on the statistical model of Q-shift dual-tree complex wavelet coefficients [J]. Journal of Optoelectronics Laser, 2009, 20(9):1252-1257.
- [19] KENNEL P, BORIANNE P, SUBSOL G. An automated method for tree-ring delineation based on active contours guided by DT-CWT complex coefficients in photographic images: Application to Abies alba, wood slice images [J]. Computers & Electronics in Agriculture, 2015, 118(C):204-214.
- [20] 李伟芳, 滕奇志, 汪华章. 基于分块直方图和共生矩阵的图像检索方法 [J]. 电子测量技术, 2009, 32(1):1-4.
- LI W F, TENG Q ZH, WANG H ZH. Image retrieval based on blocked histograms and co-occurrence matrix [J]. Electronic Measurement Technology, 2009, 32(1):1-4.

作者简介



许良凤, 1970 年出生, 副教授, 硕士生导师, 主要研究方向为多媒体通信、图像处理等。

E-mail: xulfcjn@hfut.edu.cn

Xu Liangfeng was born in 1970, associate professor, M. Sc. tutor. The main research interests include multimedia communication, and image processing, etc.



王家勇, 1992 年出生, 合肥工业大学在读研究生, 主要研究方向为数字图像处理。

E-mail: wangjiayongbfwq@163.com

Wang Jiayong was born in 1992, M. Sc. candidate in Hefei University of technology. The main research interest is digital image processing.



胡敏, 1967 年出生, 教授, 硕士生导师, 主要研究方向为计算机视觉、数字图像处理等。

E-mail: jsjxhumin@hfut.edu.cn

Hu Min born in 1967, professor, M. Sc. tutor. The main research interests include computer graphics, and digital image processing, etc.