

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2014)09-1297-09

论文引用格式: Wang X H, Jin C, Ren F J, Hu M. Research on facial expression recognition based on pyramid Weber local descriptor and the Dempster-Shafer theory of evidence[J]. Journal of Image and Graphics, 2014, 19(9): 1297-1305. [王晓华, 金超, 任福继, 胡敏. Dempster-Shafer 证据融合金字塔韦伯局部特征的表情识别[J]. 中国图象图形学报, 2014, 19(9): 1297-1305.] [DOI:10. 11834/jig. 20140906]

Dempster-Shafer 证据融合金字塔 韦伯局部特征的表情识别

王晓华¹, 金超¹, 任福继^{1,2}, 胡敏¹

1. 合肥工业大学计算机与信息学院 情感计算与先进智能机器安徽省重点实验室, 合肥 230009;
2. 德岛大学 先端技术科学教育部, 日本 德岛 7708502

摘要: 目的 针对韦伯局部特征(WLD)在局部细节表征上的局限性,提出一种基于金字塔韦伯局部特征(PWLD)的人脸表情识别方法。方法 首先对人脸图像进行预处理,分割出表情显著区域;然后根据显著区域大小进行分层并加以分块,对每一层的不同块提取 PWLD 特征,将测试集与训练集的 PWLD 直方图的卡方距离归一化作为独立证据来构造基本概率分配(BPA);最后利用 D-S(Dempster-Shafer)证据理论联合规则得到融合后的 BPA,进而得出识别结果。结果 在 JAFFE 库与 Cohn-Kanade 库上进行交叉验证,平均识别率分别为 95.81% 和 97.47%。结论 PWLD 特征在局部细节表征上比 WLD 特征更具有优势。与目前典型的方法进行对比实验,验证该方法的有效性和容错性,同时证明该方法在噪声和光照条件下具有一定的鲁棒性。

关键词: 表情识别;韦伯局部特征(WLD);金字塔韦伯局部特征;Dempster-Shafer 证据理论

Research on facial expression recognition based on pyramid Weber local descriptor and the Dempster-Shafer theory of evidence

Wang Xiaohua¹, Jin Chao¹, Ren Fuji^{1,2}, Hu Min¹

1. School of Computer and Information of Hefei University of Technology, Anhui Province Key Laboratory of Affective Computing and Advanced Intelligent Machine, Hefei 230009, China;
2. University of Tokushima, Graduate School of Advanced Technology & Science, Tokushima 7708502, Japan

Abstract: **Objective** Feature extraction is the most critical step in pattern recognition, and facial expression recognition is no exception. Weber Local Descriptor (WLD) is a method that can effectively extract texture information from images and has the advantages of being consistent with human perception of human beings and being insensitive to noise and non-monotonic illumination variations. However, WLD has some limitations in the feature representation of local details. To overcome these limitations, a facial expression recognition method based on Pyramid WLD (PWLD) is proposed in this study.

Method First, facial images are preprocessed. This step includes the detection of faces from facial expression databases and normalization. The salient regions of segment 2 that have significant contributions to facial expression recognition from images are also preprocessed. One of these salient regions is that which includes the eyes and eye brows, while another is that with the mouth. The sizes of salient regions differ, and these regions contain different information. Thus, we stratify these

收稿日期:2014-02-14;修回日期:2014-04-28

基金项目:国家自然科学基金青年基金项目(61300119);国家高技术研究发展计划(863)基金项目(2012AA011103);安徽省科技攻关基金项目(1206c0805039)

第一作者简介:王晓华(1976—),女,副教授,硕士生导师,2005 年于中国科学院研究生院获工学博士学位,主要研究方向为数字图像处理、情感计算等。E-mail: xh_wang@hfut.edu.cn

salient regions and divide each layer into different blocks. The PWLD features of each block in each layer are then extracted and cascaded to represent the global and local features of a salient region reasonably, with some parameter adjustments. Second, we compute for the Chi-square distance of the PWLD histograms in both the testing and training sets. We then choose the minimum distance in every category of expressions and normalize this distance to construct the Basic Probability Assignment (BPA) as independent evidence. To create the BPA, we use curve fitting in numerical analysis by simulating several sets of data. Finally, fusion BPA is obtained by using the Dempster-Shafer rule, and the results are further obtained by employing the decision-making and judgment of Dempster-Shafer theory of evidence. **Result** By fusing the PWLD features of the two different salient regions with Dempster-Shafer theory of evidence, we can overcome the limitations of a single regional feature and acquire more reliable and accurate results. We conduct some cross-validation experiments on the JAFFE and Cohn-Kanade facial expression databases, and the average recognition rates reach up to 95.81% and 97.47%, respectively. In addition, we perform some experiments with other algorithms, such as LBP, LDP, and Gabor; we also conduct some comparative experiments that combine the PWLD with different classifiers, such as 1-NN and SVM. **Conclusion** The WLD, which is known as a robust image descriptor, can well extract the texture information of images. Moreover, the PWLD can accurately describe the local details, which have more advantages than the WLD features. The comparative results of some typical method verify the effectiveness and fault tolerance of the proposed method. The proposed method has certain robustness under simultaneous noise and light conditions.

Key words: facial expression recognition; Weber local descriptor (WLD); pyramid Weber local descriptor (PWLD); Dempster-Shafer (D-S) theory of evidence

0 引言

在日常生活中,交流对于完成各项事务都是至关重要的,而面部表情是非口头沟通中最有效的形式,它直观地提供了关于情绪状态、心态和意图等信息^[1-3]。近些年来,随着人机交互的发展,人脸表情识别(FER)已经成为模式识别领域研究热点之一^[4-5]。人脸表情识别主要分为3个步骤:目标检测与预处理、特征提取、分类识别,而人脸特征的提取被认为是自动人脸表情识别的最关键部分^[6]。目前,很多特征提取方法被提出来,大致可以分为:基于几何特征的方法和基于纹理特征的方法,如Gabor小波^[7]、局部二值模式(LBP)^[8]、基于活动外观模型(AAM)^[9]等方法,实际的识别算法中常常采用多特征融合的方法来解决人脸表情识别中光照、姿态、噪声等问题。

基于Gabor小波的方法提取不同尺度和不同方向的信息,对于不同表情具有很好的区分性,但由于其过程耗时且特征维数较高,很难将其应用到实时系统中;LBP提取的特征常被用在纹理信息的描述,其计算简单,同时具有一定的灰度不变性和尺度不变性^[8],但是LBP对于图像中光照和噪声敏感。Jabid等人^[6]分析了LBP的缺点,提出了计算效率更高的局部差分模式(LDP),并且认为LDP对光照和噪声鲁棒。由于每个像素点都要计算8个方向的

Kirsch算子的值,同时令其中 k 个绝对值最大的对应的位置为1,其他位的值置为0,虽然可以降低维数,但是相应地增加了计算复杂度并且降低了对纹理的表征能力。Su等人^[10]融合了Gabor和LBP两组特征在表情识别上取得了很高的识别率,但是其维数很大,特征提取过程也非常耗时。陈杰等人^[11]提出一种简单、具有很强描述能力并且鲁棒的局部描述子,名为韦伯局部特征(WLD),它是基于心理学中的韦伯定律提出来的,通过模拟人感知周围环境来提取特征的,所以它在图像纹理提取方面非常有效,并且对光照和噪声鲁棒。Yu等人^[12]将WLD特征用于网络图像中识别人脸表情,先从大量的网络图像中检测到人脸,再通过主动学习去除有噪声的图像,得到一系列的人脸图像,然后再对人脸图像分块加权,提取不同尺度的WLD特征,取得了不错的效果。但是由于权值和分块固定,具有局限性,不能更好地提取局部细节特征。

针对传统的WLD对于局部描述不强的问题,本文提出了一种金字塔韦伯局部特征(PWLD),使用金字塔来提取不同大小不同区域的信息,最后结合所有块的PWLD直方图序列可以较合理地描述人脸图像的全局和局部特征。首先提取一幅图像的WLD特征,再将图像进行分层,每一层为上一层的各块按宽和高等分成更小的块,并通过调整WLD中的参数对每一个块提取WLD特征,最后按照一定权值将所有的WLD特征合并成最终的

PWLD 特征。对于来自不同方面的特征, 由于 Dempster-Shafer (D-S) 证据理论是通过采用区间估计的方法描述不确定的信息, 在区分不确定与不知道方面具有很大的灵活性, 所以本文采用 D-S 证据理论数据融合不同方面的信息, 可以克服单一特征的局限性, 从而得到更为可靠、准确的结果。在提取人脸图像中两个显著区域的 PWLD 特征后, 计算测试样本与每个训练样本特征间的距离, 通过数值分析中的曲线拟合构造出基本概率分配 (BPA), 最后通过 D-S 证据理论融合规则得出最后的识别结果。通过在不同表情库上进行实验并与其他方法比较, 其结果表明, 本文方法在识别率和识别时间上更有优越性。

1 PWLD 的特征提取方法

1.1 韦伯局部特征

韦伯定律表明能够被感知的刺激强度 ΔI 与原刺激值 I 大小的比值是一个常数。 $\Delta I/I$ 被称为韦伯比。这意味着刺激强度变化的比值反映了人类感知刺激的幅度。受到该定律的启发, 陈杰等人^[11]提出了 WLD 用以描述图片的纹理信息, 并认为由 WLD 提取的特征更符合人类感知, 对光照和噪声具有很好的鲁棒性。

WLD 由两个部分组成: 差分激励 $\xi(x_c)$ 和方向 $\theta(x_c)$ 。由于 WLD 希望找到图像中较为显著的变化来模拟人的感知, 所以 WLD 使用一个像素 x_c 和它邻域像素之间的强度差分的比值作为该像素的变化。具体来说, 图像中任一像素 x_c 的差分激励为

$$\xi(x_c) = \arctan \left[\sum_{i=0}^{p-1} \left(\frac{x_i - x_c}{x_c} \right) \right] \quad (1)$$

式中, x_i 表示当前像素 x_c 的第 i 个邻域像素, p 是邻域像素的个数。WLD 另一个组成部分是方向 $\theta(x_c)$, 它指的是像素 x_c 的梯度方向, 即

$$\theta(x_c) = \arctan \left(\frac{v_s^a}{v_s^b} \right) \quad (2)$$

式中, v_s^a 和 v_s^b 分别由滤波窗口函数 f_{00} 和 f_{01} 算出的, 滤波窗口函数为

$$f_{00} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & +1 & 0 \end{bmatrix}, f_{01} = \begin{bmatrix} 0 & 0 & 0 \\ +1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \quad (3)$$

为了简化起见, 将 θ 量化到 T 个主要的方向。具体量化的方程为

$$\Phi_t = f_q(\theta') = \frac{2t}{T} \pi \quad (4)$$

$$t = \text{mod} \left(\left\lfloor \frac{\theta'}{2\pi/T} \right\rfloor, T \right)$$

式中, θ' 是由 $f: \theta \mapsto \theta'$ 计算得来, 即

$$\theta' = \arctan 2(v_s^a, v_s^b) + \pi$$

$$\arctan 2(v_s^a, v_s^b) = \begin{cases} \theta & v_s^a > 0 \quad \& \quad v_s^b > 0 \\ \pi + \theta & v_s^a > 0 \quad \& \quad v_s^b < 0 \\ \theta - \pi & v_s^a < 0 \quad \& \quad v_s^b < 0 \\ \theta & v_s^a < 0 \quad \& \quad v_s^b > 0 \end{cases} \quad (5)$$

根据式 (1) (5), 计算出差分激励 $\xi(x_c)$ 和方向 $\theta(x_c)$ 后, 就可以得出 2 维 WLD 直方图 $\{WLD(\xi_j, \Phi_t), j=0, 1, \dots, N-1; t=0, 1, \dots, T-1\}$, 其中, N 是图像的维度。为了使 WLD 特征区分能力更强, 2 维 WLD 直方图被进一步转化为 1 维直方图 H 。

1.2 金字塔韦伯局部特征 (PWLD)

WLD 可以很好地提取纹理信息, 具有较高的区分度, 但是对于局部纹理并不能很好地表征, 本文提出一种 PWLD。其特征表示是通过逐级分割图像, 再将每一级乘以一个归一化系数 α , 最后综合每一级的每一块直方图, 形成最终的 PWLD 直方图, 使得 PWLD 特征对局部特征的描述更加精确化。

图 1(a) 是原图像, 图 1(b) 为第 0 级图像分割及其对应的该级 PWLD 直方图, 第 0 级的 PWLD 直方图的提取即是对整幅图像 PWLD 直方图的提取。根据图像大小和对局部特征提取的需要, 可以将图像分成 L 级, 并分别在每一块上提取 PWLD 特征。

PWLD 直方图的提取过程如下: 由式 (1) (5) 求出图像的差分激励 $\xi(x_c)$ 和方向 $\theta(x_c)$, 可以联合得到图像的 2 维直方图, 在这里, 为了降低噪声的影响, 采用给韦伯比 $\Delta I/I$ 设置一个上下界限 $K_{\min}$, $K_{\max}$ 。在这个 2 维直方图里, 每一列对应着一个方向 $\Phi_t (t=0, \dots, T-1)$, 每一行对应着不同的差分激励, 所以每一个主方向 Φ_t , 都相当于一个 1 维直方图 $H(t)$ 。为了使该 2 维直方图具有更好的区分能力, 再将 $H(t)$ 平均分成 M 小段, 所有的直方图分段 $H_{m,t} (m=0, 1, \dots, M-1)$ 形成了直方图矩阵。最后再将每一个子直方图 $H_{m,t}$ 细分成 S 个子区间, 即 $H_{m,t} = \{h_{m,t,s}, s=0, 1, \dots, S-1\}$, 其中 $h_{m,t,s}$ 为

$$h_{m,t,s} = \sum_j \delta(S_j = s)$$

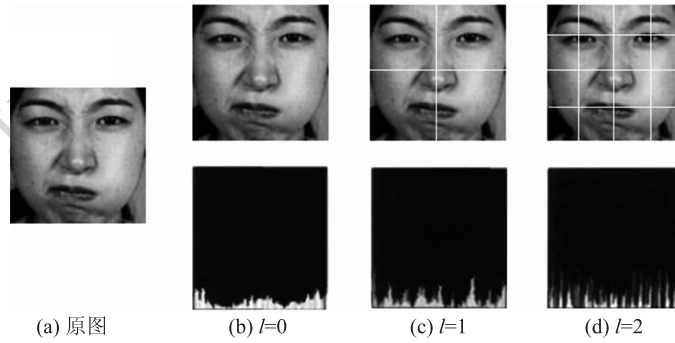


图1 3层 PWLD 特征描述

Fig. 1 Three layers of characterization of PWLD

$$S_j = \left\lfloor \frac{\xi_j - \eta_{m,l}}{(\eta_{m,u} - \eta_{m,l})/S} \right\rfloor + \frac{1}{2} \quad (6)$$

$\delta(\cdot)$ 是 Kronecker 差分方程, 即

$$\delta(X) = \begin{cases} 1 & X \text{ 为真} \\ 0 & \text{其他} \end{cases} \quad (7)$$

$\eta_{m,l}$ 、 $\eta_{m,u}$ 是 M 个子段的下界 $\eta_{m,l} = (m/M - 1/2)\pi$ 和上界 $\eta_{m,u} = [(m+1)/M - 1/2]\pi$ 。

由于 PWLD 将图像分成 L 级, 所以对于不同级数, 设置对应的 $M_i, T_i, S_i (i=0, \dots, L-1)$, 进一步细致地表征纹理信息。由图 1 可以看出, 随着分割层数的增加, 对于图像描述的刻画越来越局部化和精细化, 提取的 PWLD 特征再进行归一化, 将第 l 级直方图乘以归一化系数 $\alpha = 1/4^l$ 。最终的 PWLD 特征维数为 $\sum_{i=0}^{L-1} 4^i M_i T_i S_i$, 时间复杂度 $O_{\text{PWLD}} = Cmn$, 其中 $m \times n$ 是图像的大小, C 是一个常量, 表示每个像素进行加、除、求反正切以及滤波等运算。

2 基于 PWLD 和 D-S 理论的表情识别

2.1 D-S 证据理论

D-S 证据理论^[13]最初是由 Dempster 提出, Shafter 对其进行完善的, 又被称为 D-S 理论。它将两个或者多个证据体的基本信任函数通过 D-S 合成规则融合成一个新的基本信任函数, 作为最终的决策依据, 因此可以获得更高的识别率和可靠性。

D-S 理论用识别框架 $\Theta = \{S_1, S_2, \dots, S_n\}$ 来表示对于一个判决问题可以认识到的所有结果的集合, 并定义了一个集函数 $m: 2^\Theta \rightarrow [0, 1]$ 为识别框架 Θ 上的基本概率分配函数 (BPA), 满足条件

$$\begin{cases} m(\emptyset) = 0 \\ \sum_{A \subseteq \Theta} m(A) = 1 \end{cases} \quad (8)$$

$m(A)$ 反映了对 A 本身的信任度, $m(\emptyset)$ 表示不确定性。对于两个独立的证据, 假设 m_1, m_2 是两个对应的基本概率分配函数, 焦点分别是 $A_1, \dots, A_M$ 和 $B_1, \dots, B_M$, 则根据 D-S 证据理论的合成法则得到合成后的基本概率分配函数 $m(C)$, 具体计算为

$$\begin{cases} m(\emptyset) = 0 \\ m(C) = \sum_{A_i \cap B_j = C} \frac{m_1(A_i)m_2(B_j)}{1-K}, C \neq \emptyset \end{cases} \quad (9)$$

式中, $K = \sum_{A \cap B \neq \emptyset} m_1(A_i)m_2(B_j)$, 反映了每个证据之间的冲突程度。

2.2 基本概率分配函数 (BPA) 的构造

在表情识别中, 识别框架 Θ 是一个 N 元组, 即有 N 类表情, 设 $\Theta = \{\mu_1, \mu_2, \dots, \mu_N\}$, 在本文方法中, 证据来自于人脸中两个显著区域 $i (i=1, 2)$ 。在提取完每个训练样本和要测试的样本的 PWLD 特征后, 计算测试样本与训练样本库中所有样本的卡方距离, 取每一类表情中最小的卡方距离 $d_{i,j}$, i 指的是显著区域, j 是表情的类别。由于 $d_{i,j}$ 是 $(0, +\infty)$ 上任意实数, 所以先对所有的 $d_{i,j}$ 进行归一化, 即

$$d_{i,j} = \frac{d_{i,j}}{\sum_{j=1}^N d_{i,j}} \quad (10)$$

经过归一化后, 所有的 $d_{i,j} \in [0, 1]$, 将它们由小到大排序。当距离越接近 0 时, 基本概率分配值越接近于 1, 为了增大 $d_{i,j}$ 中前几个值的区分度, 使之有利于决策判断, 同时降低后几个非主要因素的影响, 下面将构造一个曲线来完成距离到基本概率分配值的映射。该曲线应该具有如下特点: 存在在 $x = x_0$ 处斜率为 -1, 在 $(0, x_0)$ 区间内所有点的斜率小于 -1, 而在 $(x_0, 1)$ 区间内曲线则相对平缓。经过反复试验, x_0 在 $(0.35, 0.4)$ 区间内可以取得最佳

的效果,本文取 $x_0 = 0.37$ 。根据曲线的特点,通过在坐标轴上模拟几组数据,再运用数值分析中的曲线拟合,并作一些参数调整,构造出指数函数

$$y = e^{-13(x+0.06)^2} + 0.05 \quad (11)$$

具体曲线如图2所示。

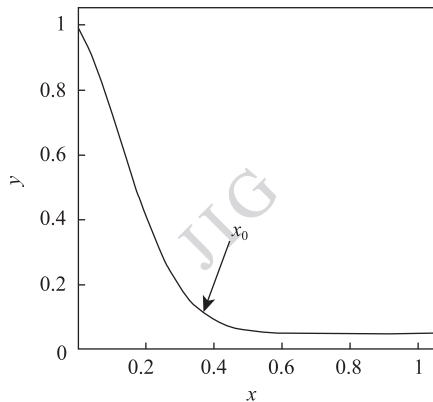


图2 函数的曲线

Fig.2 Curve of the function

根据构造好的函数表达形式和式(8)两个条件,定义证据 i 对类别 μ_j 的 BPA 为

$$m_i(\mu_j) = \beta \times (e^{-13(d_{i,j}+0.06)^2} + 0.05) \quad (12)$$

式中, β 为调节系数,使得 $\sum_{j=1}^N m_i(\mu_j) < 1$, 如果 β 为 0, 说明目标完全被噪声淹没, 不确定性最大。根据定义, 将不确定性分配函数构造为

$$m_i(\theta) = 1 - \sum_{j=1}^N m_i(\mu_j) \quad (13)$$

2.3 联合规则及决策判断

设 $\mu_i (i = 1, 2, \dots, N)$ 为表情类别, N 是表情的类别数, A_r 为目标类别。经过式(10)(12)(13)求得融合后的信任度 $m(A_r)$ 和不确定性 $m(\theta)$ 必须满足以下规则:

- 1) $m(A_r)$ 应具有最大的基本概率分配值;
- 2) 目标类别 $m(A_r)$ 与其他类基本概率分配值应大于某个门限 ε_1 ;
- 3) 不确定性 $m(\theta)$ 必须小于某个门限 ε_2 ;
- 4) 目标类别 $m(A_r)$ 必须大于不确定性 $m(\theta)$ 。

2.4 基于 PWLD 和 D-S 理论的表情识别方法

本文方法主要分为 3 个阶段: 人脸检测与预处理、特征提取、分类识别。系统流程图如图3所示。

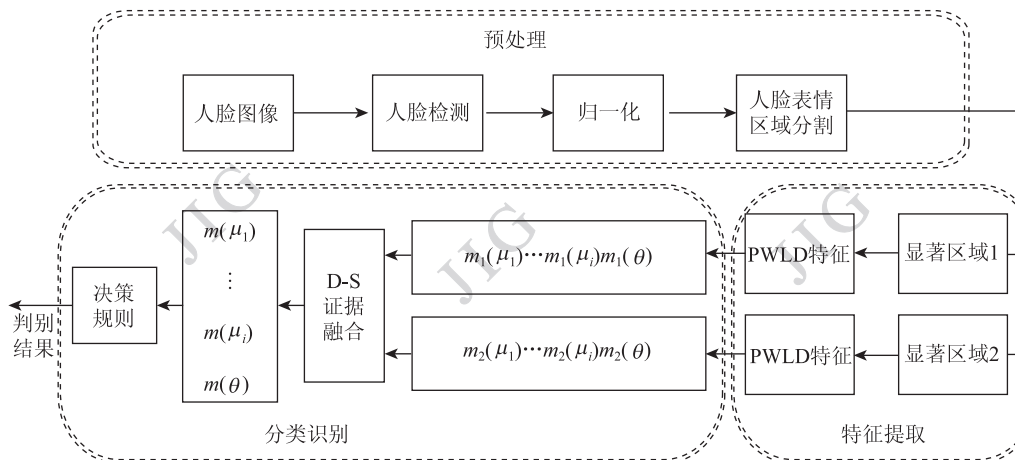


图3 人脸表情识别系统流程图

Fig.3 The flow chart of facial expression recognition system

具体步骤如下:

1) 预处理。先用 AdaBoost^[14] 人脸检测算法检测人脸, 由于人的五官具有一个统计规律, 再用双向灰度积分投影法对检测出的人脸进行眼睛的定位^[15], 然后对图像进行角度、尺寸的归一化处理。本文将所有图像归一化为 128×128 像素。由于人脸中眼睛、眉毛和嘴巴对于表情识别贡献最大^[16], 所以分割出两个显著区域: 显著区域1(眉毛、眼睛), 显著区域2(嘴巴)。

2) 特征提取。由步骤1)得到的两个显著区域, 分别对它们提取 PWLD 特征。由于显著区域1包括眉毛和眼睛, 含很多信息, 所以将图像分成3层, 分别提取每个块的 PWLD 特征, 在系统中, 3级 PWLD 特征中的 (M_i, T_i, S_i) 分别取 $(6, 8, 10)$, $(3, 8, 5)$, $(3, 8, 2)$ 。对于显著区域2, 由于区域大小的限制, 将图像分成两层, (M_i, T_i, S_i) 和显著区域1中的 $l = 0$ 级和 $l = 1$ 级参数一样。为了进一步抑制图像中噪声的影响, 设置韦伯比的上下界分别为 $K_{\min} = 0.1$,

$K_{\max} = 3$ 。

3)分类识别。构造基本概率分配函数 m_1, m_2 ，并根据 D-S 证据融合得到融合后的基本概率分配函数 m ，再由决策规则，得出判别结果。

3 实验结果与分析

为了验证本文方法的有效性，采用 JAFFE 人脸

表情库^[17]和 Cohn-Kanade (C-K) 人脸表情库^[18]进行实验。JAFFE 人脸表情库是由 10 位日本女性的 213 幅人脸图像构成，包含有 7 种基本表情：生气、厌恶、恐惧、高兴、中性、悲伤、吃惊。每个人包含 2~4 幅不同表情的图像。Cohn-Kanade 人脸表情库包含了 210 个对象的 6 种表情序列：悲伤、高兴、吃惊、恐惧、生气、厌恶。两个库归一化后的部分图像如图 4 所示。

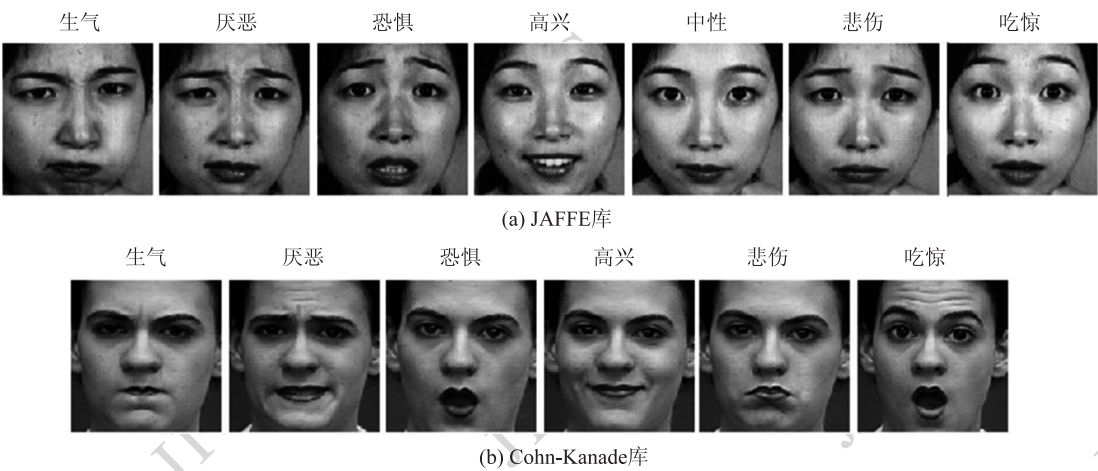


图 4 JAFFE 和 Cohn-Kanade 人脸库中的部分样本

Fig. 4 Some samples from JAFFE and Cohn-Kanade face database

3.1 本文方法的实验结果及分析

实验中，利用本文方法在 JAFFE 人脸库上进行了 10 次实验，每次随机从库中选取每人每种表情的 1~2 幅图像作为训练，剩下的作为测试，并保持每种表情的测试样本数为 15 幅。在 Cohn-

Kanade 人脸库里，选取 30 个对象的 6 种表情，每种表情选取 4 幅图像，任意取其中 2 幅进行训练，剩下的进行测试。同样也在 Cohn-Kanade 人脸库上进行了 10 次实验。最终的实验结果如表 1 所示。

表 1 在 JAFFE 人脸库和 Cohn-Kanade 人脸库的 10 次实验结果

Table 1 Experiment results of 10 times on JAFFE face database and Cohn-Kanade face database

表情	JAFFE 人脸库			Cohn-Kanade 人脸库		
	测试样本数	平均识别率/%	标准差	测试样本数	平均识别率/%	标准差
生气	15	96	0.032 7	60	96.67	0.010 5
厌恶	15	90.67	0.032 7	60	96.5	0.011 7
恐惧	15	95.33	0.030 6	60	97.33	0.011 1
高兴	15	99.33	0.02	60	98.5	0.011 7
悲伤	15	97.33	0.032 7	60	96.33	0.01
吃惊	15	98.67	0.026 7	60	99.5	0.007 6
中性	15	93.33	0.029 8			
平均		95.81			97.47	

由表 1 中的数据可以看出，本文方法还是存在一些误识别的情况，图 5 列出了部分误识别的图像，

原因主要有以下几点：

1) 表情之间存在微表情，相似度很高，如图 5

(b)(d), 属于微表情的误识别;

2) 有些表情在两个显著区域的细节变化很相似, 如图 5(a)(c), 造成融合时不确定性变大;

3) 有的表情里存在比较大的姿态变化。



图 5 部分误识别图像

Fig. 5 Part of the misrecognized images

3.2 不同算法的性能比较

同样在 JAFFE 人脸库和 Cohn-Kanade 人脸库上与 Gabor 小波、LBP 和 LDP 等方法比较他们的识别率和识别时间, 并将 PWLD 特征结合不同的分类器进行比较, 这里与其他两种分类器比较时采用的是投票机制, 分别提取人脸图像两个显著区域的 PWLD 特征, 接着将测试样本和训练样本的特征送入到分类器中可以得到两个结果, 如果结果一致则输出识别结果, 结果不一致则将两组特征级联起来作为该人脸的第 3 组特征, 送入分类器后可以得到 3 个结果, 票数多的为最终结果, 如果都不一样, 则认为识别失败。为了证明本文 PWLD 特征比原有的 WLD 特征具有更好的表征局部信息能力, 同样也用 WLD 特征进行对比实验。实验在 Windows XP 系统上 (CPU 2.19 GHz, 内存 1.99 G), 用 VC 6.0 实现的。实验结果如表 2—表 4 所示。其中 SVM 采用的是多项式核函数, $d=5$, 实验中涉及的 Gabor 均采用 5 尺度、8 方向。

表 2 表明, 用本文 PWLD 特征与 D-S 证据理论结合可以获得更高识别率, 尤其对于一些微表情有较高的区分度, 表 5 为 D-S 方法的数据融合结果, 选取两类容易出现误识别的表情作为应用举例, $\mu_i (i=1, 2, \dots, 7)$ 分别表示生气、厌恶、恐惧、高兴、中性、悲伤、吃惊, 从表 5 中的数据可以看出, 不同显著区域得出的结果不一致, 但是由融合结果可以看出本文构造的基本概率分配函数的正确性以及 D-S

证据理论在人脸表情识别的有效性和容错性。

由表 3 和表 4 结果反映出: 相比于其他方法, 本文方法有更高的识别率、较低的维数和识别时间。虽然识别率还不能达到 Gabor + LBP^[10] 的效果, 但是在特征维数和识别时间上远低于它。通过与传统的 WLD 特征的比较, 虽然在识别时间和特征维数上有些增加, 但是本文方法的识别率有很大的提高。验证了用本文方法提取局部信息比传统的 WLD 更加有效。

表 2 不同分类器平均识别率的比较

Table 2 Comparison of average recognition rates of different classifiers

特征 + 分类器	JAFFE/ %	Cohn-Kanade/ %
PWLD + 1-NN	92.06	92.31
PWLD + SVM	93.97	95.09
PWLD + D-S	95.81	97.47

表 3 不同方法平均识别率的比较

Table 3 Comparison of average recognition rates of different algorithms

算法	相关参数	平均识别率/ %	
		JAFFE	C-K
LBP ^[8]	3 × 3 窗口	90.16	92.59
LDP ^[6]	7 × 6 个子块, $k=3$	90.79	97.04
Gabor 小波 ^[7]	默认	91.75	92.69
Gabor + LBP ^[10]	默认	97.14	98.15
WLD	$M=6, T=8$ $S=20$	91.43	93.06
PWLD	默认	95.81	97.47

表 4 不同算法平均识别时间的比较

Table 4 Comparison of average recognition time of different algorithms

算法	特征维数	平均识别时间/ms	
		JAFFE	C-K
LBP ^[8]	10 710	25.7	26.1
LDP ^[6]	2 352	66.8	68.6
Gabor 小波 ^[7]	368 640	2 702	2 751
Gabor + LBP ^[10]	14 336	34 050	34 080
WLD	960	111.4	115.9
PWLD	1 728	437.5	440.7

表 5 D-S 方法的数据融合结果

Table 5 Results of data fusion using the Dempster-Shafer theory of evidence

表情图像	显著区域	基本概率分配值								识别结果
		$m(\mu_1)$	$m(\mu_2)$	$m(\mu_3)$	$m(\mu_4)$	$m(\mu_5)$	$m(\mu_6)$	$m(\mu_7)$	$m(\theta)$	
厌恶	1	0.140 1	0.163 2	0.132 4	0.137	0.137 2	0.137 3	0.105 2	0.047 6	厌恶
	2	0.131 1	0.128	0.131 5	0.123 5	0.137	0.15	0.150 5	0.048 4	不确定
	融合	0.141 1	0.156 7	0.135 2	0.132 3	0.143 6	0.154 5	0.126 3	0.010 3	厌恶
中性	1	0.093 2	0.094 2	0.143 1	0.167 2	0.160 7	0.127 2	0.170 1	0.044 3	吃惊
	2	0.124 6	0.134 2	0.156 6	0.122 5	0.173 3	0.136 8	0.104 6	0.047 4	中性
	融合	0.098	0.104 9	0.164 3	0.154	0.196 3	0.134 2	0.138 7	0.009 6	中性

4 结 论

本文基于 PWLD 和 D-S 证据理论的人脸表情识别方法,与目前的主要算法相比,不管在平均识别率上还是平均识别时间上都有一定的优越性,大量实验结果表明:

1)PWLD 特征比原始 WLD 特征具有更好的对局部信息表征的能力,对于图像局部纹理描述地更加精确,可以很好地提取全局和局部特征;

2)PWLD 特征同时还保留了原始 WLD 特征对于噪声和光照变化非常鲁棒的优点,由于其简单并且有效,可以将其用于实时系统;

3)通过 D-S 证据理论将来自不同显著区域的 PWLD 特征进行融合,并运用数值分析中的曲线拟合构造基本概率分配函数(BPA),使得本方法在人脸表情识别中取得了不错的识别效果,同时也保证了本方法的有效性以及容错性。

后续工作将在以下两个方面进行改进:

1)构造更好的基本概率分配函数(BPA)和不确定性分配函数进一步降低系统的误识别率。

2)研究更好的预处理和特征提取方法解决人脸中的姿态和遮挡等问题。

参考文献 (References)

[1] Dhall A, Goecke R, Lucey S, et al. Static facial expression analysis in tough conditions: data, evaluation protocol and benchmark[C]//Proceedings of IEEE International Conference on Computer Vision Workshops. Los Alamitos: IEEE Computer Society Press, 2011:2106-2112.

[2] Fasel B, Luetttin J. Automatic facial expression analysis: a sur-

vey [J]. Pattern Recognition, 2003, 36(1):259-275.

[3] Xue Y L, Mao X, Guo Y, et al. The research advance of facial expression recognition in human computer interaction [J]. Journal of Image and Graphics, 2009, 14(5):764-772. [薛雨丽, 毛峡,郭叶,等. 人机交互中的人脸表情识别研究进展[J]. 中国图象图形学报, 2009, 14(5):764-772.] [DOI: 10.11834/jig.20090503]

[4] Valstar M F, Jiang B H, Mehu M,et al. The first facial expression recognition and analysis challenge [C]//Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition. Los Alamitos: IEEE Computer Society Press, 2011: 921-926.

[5] Gehrig T, Ekenel H K. Facial action unit detection using kernel partial least squares [C]//Proceedings of IEEE International Conference on Computer Vision Workshops. Los Alamitos: IEEE Computer Society Press, 2011: 2092-2099.

[6] Jabid T, Kabir M H, Chae O. Facial expression recognition using local directional pattern [C]//Proceedings of IEEE International Conference on Image Processing. Los Alamitos: IEEE Computer Society Press, 2010: 1605-1608.

[7] Liu W F, Wang Z F. Facial expression recognition based on fusion of multiple gabor features [C]//Proceedings of the 18th International Conference on Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2006: 536-539.

[8] Shan C F, Gong S G, McOwan P W. Facial expression recognition based on local binary patterns: a comprehensive study [J]. Image and Vision Computing, 2009, 27(6):803-816.

[9] Zhou M C, Lin L, Sun Jian, et al. AAM based tracking with temporal matching and face segmentation [C]//Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2010: 701-708.

[10] Sun X H, Xu H X, Zhao C X, et al. Facial expression recognition based on histogram sequence of local Gabor binary patterns [C]//Proceedings of IEEE Conference on Cybernetics and Intelligent Systems. Los Alamitos: IEEE Computer Society Press, 2008: 158-163.

- [11] Chen J, Shan S, He C, et al. WLD: a robust local image descriptor [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(9):1705-1720.
- [12] Yu K M, Wang Z Y, Zhuo L, et al. Learning realistic facial expressions from web images [J]. *Pattern Recognition*, 2013, 46(8):2144-2155.
- [13] Dempster A P. Upper and lower probabilities induced by multi-valued mapping [J]. *Annals of Mathematical Statistics*, 1967, 38:325-339.
- [14] Viola P, Jones M. Robust real-time face detection [J]. *International Journal of Computer Vision*, 2004, 57(2): 137-154.
- [15] Dai J W, Liu D, Su J B. Rapid eye localization based on projection peak [J]. *PR&AI*, 2009, 22(4):605-609. [戴景文, 刘丹, 苏剑波. 基于投影峰的眼睛快速定位方法[J]. 模式识别与人工智能, 2009, 22(4):605-609.]
- [16] Wei R, Jiang L, Tao L M. Facial expression recognition system based on multiple feature integration [J]. *Journal of Image and Graphics*, 2009, 14(5):792-800. [魏冉, 姜莉, 陶霖密. 融合人脸多特征信息表情识别系统[J]. 中国图象图形学报, 2009, 14(5):792-800.] [DOI: 10.11834/jig.20090506]
- [17] Lyons M J, Budynek J, Akamatsu S. Automatic classification of single facial images [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999, 21(12):1357-1362.
- [18] Kanade T, Cohn J F, Tian Y. Comprehensive database for facial expression analysis[C]//*Proceedings of the 4th IEEE International Conference on Face and Gesture Recognition*. Los Alamitos: IEEE Computer Society Press, 2000: 46-53.