

DOI: 10.13382/j.jemi.B2003279

基于区域 NSBP 特征的加权证据融合表情识别*

刘娟^{1,2} 胡敏² 黄忠^{1,2}(1. 安庆师范大学 电子工程与智能制造学院 安庆 246133; 2. 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 230601)

摘要:为了提取鲁棒性强的人脸纹理特征并提高区域特征决策融合的性能,提出一种基于邻近平滑二值模式(neighbor smooth binary pattern, NSBP)特征描述子和加权证据融合(weighted evidence fusion, WEF)的表情识别新方法。首先,提出了一种 NSBP 描述子,通过判定水平、垂直及对角线方向上的“中心”像素点灰度值是否在各梯度方向上两邻域的灰度值范围内来对图像进行编码;然后基于提取的眉毛、眼睛和嘴巴区域的 NSBP 纹理特征来构造证据的初始基本概率分配(basic probability assignment, BPA);最后针对登普斯特-谢弗(Dempster-Shafer, D-S)证据理论在证据之间存在冲突时进行融合的不足,提出一种加权证据修正的合成方法,以完成3个区域证据的决策融合。实验结果表明,该方法在CK(Cohn-Kanade)数据库上的平均表情识别率和识别时间分别为95.25%、765 ms,与其他相关方法的比较也验证了其有效性。

关键词:表情识别;NSBP特征;加权证据融合;Dempster组合规则

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Weighted evidence fusion expression recognition based on regional NSBP features

Liu Juan^{1,2} Hu Min² Huang Zhong^{1,2}(1. School of Electronic Engineering and Intelligent Manufacturing, Anqing Normal University, Anqing 246133, China;
2. Anhui Province Key Laboratory of Affective Computing and Advanced Intelligent Machine, School of Computer and
Information, Hefei University of Technology, Hefei 230601, China)

Abstract: In order to extract robust facial features and improve the decision-level fusion of multi-regional features, a new expression recognition method based on neighbor smooth binary pattern (NSBP) feature descriptor and weighted evidence fusion (WEF) is proposed. First, a NSBP descriptor is proposed to encode the image by determining whether the gray values of the center pixels in the horizontal, vertical and diagonal directions are within the gray value range of two neighborhoods in each gradient. Then the initial basic probability assignments (BPA) of evidences are constructed based on the extracted NSBP texture features of the eyebrows, eyes, and mouth regions. Finally, aiming at the deficiency of Dempster-Shafer (D-S) evidence theory in conflict evidence fusion, a synthetic method of weighted evidence revision is proposed to realize the decision fusion of three regional evidences. Experimental results show that the recognition rate of this method on the Cohn-Kanade (CK) database is 95.25%, and the average recognition time is 765 ms, compared with other related methods, the effectiveness of this method is also verified.

Keywords: facial expression recognition; neighbour smooth binary pattern feature; weighted evidence fusion; Dempster combination rule

收稿日期: 2020-06-29 Received Date: 2020-06-29

*基金项目: 国家自然科学基金(61672202, 61702012)、安徽省自然科学基金(1908085MF195)、安徽省重点实验室开放课题项目(ACA1M180203)、中央高校基本科研业务费专项资金(PA2020GDSK0061)、安徽省高校协同创新项目(GXXT-2019-030)资助

0 引言

随着智能信息的快速发展,表情识别当作一种人机交互的方式,同时也得到了学术界的广泛关注^[1]。由于个人存在外貌和表情表现形式的不同,所提取的表情特征一般包括待测者的个人信息,从而给表情识别过程增加了难度。怎样提取具有较强鲁棒性和较高鉴别能力的特征,以及减少差异性信息的不利作用,是面部表情识别的热门研究课题^[2]。

近年来,人们已经研究了多种不同的人脸特征提取方法,重点包含基于面部变形的形状特征和基于像素信息的纹理特征。形状特征提取方法通常可分为活动形状模型(ASM)和活动外观模型(AAM)这两大类^[3-4],它们可以通过迭代求解在不同约束下来优化面部特征点提取,然而存在容易陷入局部最小值、计算复杂和计算速度不快的缺陷。本文着重于纹理特征的提取,它基于相邻区域像素周围灰度的分布和对比度描述了表情的局部细节。最常见的纹理特征提取方法包括定向梯度直方图(histogram of oriented gradient, HOG)^[5], Gabor 小波^[6-7]和局部二值模式(local binary pattern, LBP)^[8]。HOG 特征是通过计算不同梯度方向上的梯度大小分布而构造的,具有计算量较小的优点,但不具有旋转不变性和尺度不变性。Gabor 小波可以描述不同尺度、不同方向的图像信息,因此具有很高的精度,但是增加了时间和空间的复杂性,不利于实时交互。与多尺度多方向的 Gabor 特征相比,LBP 特征在特征维数和特征提取时间方面拥有优势,进而在诸如人脸表情识别之类的实时场景中应用更胜一筹。LBP 作为一种能够表示图像局部细节的算子,除了计算简单以外,还拥有灰度不变性和旋转不变性的好处。所以它被广泛地应用于机器视觉和模式识别等领域,同时出现了众多改进的算法。但是 LBP 仅将中心像素和它邻近像素点的灰度值作对比,未探讨邻域像素点灰度值之间的大小关系,这造成部分构造信息的缺失^[9]。

就表情分类而言,研究者逐步研究将面部划分为很多表情特征区域并提取其特征,同时引入了基于区域特征的数据级联^[10]和决策融合方法。数据级联方法是根据一定的规则组合不同的区域特征,从而形成统一的判别向量,并把经过降维后特征向量的主要成分输入一个或多个分类器中来完成识别。和数据级联方法相比较,区域特征决策融合方法不同的是通过区域特征获取类别度量信息,其次在决策层中将它们进行融合。鉴于登普斯特-谢弗(Dempster-Shafer, D-S)证据理论拥有可以很好地表示不确定性的优点,因此涌现出许多决策融合面部表情识别方法,如王晓华等^[11]使用来源于面部不同区域的 PWLD 特征作为证据,并采用 D-S 证据理论来实现这

些证据的合成。Shoyaib 等^[12]通过基于距离度量的 K 最近邻(KNN)特征信息来计算证据的概率分布函数,然后使用 D-S 证据理论将其融合。但是当融合的证据相互之间存在较高冲突时,传统的 D-S 证据理论可能会出现和直觉结果相反的情形,从而导致对目标的决策无效^[13]。

针对面部表情特征重点聚集在眉毛、眼睛和嘴巴的关键部位,而且基本的 LBP 算子不能很好地描绘局部纹理信息,因此提出基于加权证据融合(weighted evidence fusion, WEF)的区域邻近平滑二值模式(neighbor smooth binary pattern, NSBP)的表情识别方法。首先鉴于传统的 LBP 算子只是统计中心像素与邻域像素之间的灰度值差别,提出了一种 NSBP 的纹理特征描述算子,这种算子是通过判断水平、垂直及对角线这 3 个方向上的“中心”像素点的灰度值是不是在每个梯度方向上的两邻近像素点灰度值界限内来对图像实行编码。其次鉴于面部表情信息主要反映在局部关键区域,如眉毛、眼睛和嘴巴,由这些关键区域提取的 NSBP 纹理特征被用作证据,并且依据卡方距离来建立每个证据的初始概率赋值。最后鉴于传统 D-S 证据理论在融合冲突证据时的缺点,提出了基于 WEF 策略来完成 3 个关键区域证据的决策融合。

1 邻近平滑二值模式

1.1 LBP

LBP 运算符已成功应用于面部识别^[14]。在 3×3 域中定义的基本 LBP 特征仅只计算中心像素和附近邻域的灰度值大小关系。假设附近邻域的灰度值不小于中心像素点,那么该像素点地方记为 1;否则记为 0。通过作对比,从附近邻域内的 8 个像素点能够形成一个 8 bits 二进制编码序列。将获得的编码序列转变成对应的十进制数,然后把它用作中心像素的 LBP 特征编码值,它能够体现中心像素附近区域的纹理信息。LBP 的编码运算公式如下:

$$LBP = \sum_{p=0}^{P-1} f(g_p - g_c) 2^p \quad (1)$$

$$f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)$$

式中: g_c 表示中心像素点的灰度值; g_p 表示附近邻域内像素点的灰度值; P 表示附近邻域内像素点的数目; $f(x)$ 表示符号函数。基本 LBP 算子的计算过程如图 1 所示。

1.2 NSBP

LBP 运算符仅比较中心像素和它附近邻域像素点灰度值大小,并没有考虑附近邻域像素点之间的灰度值关系。本文提出了 NSBP,该特征方法是通过判定水平、垂

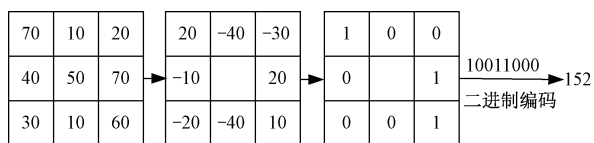


图 1 基本 LBP 算子的计算过程

Fig. 1 Calculation process of the basic LBP operator

直及对角线这 3 个方向上“中心”像素点灰度值是否在每个梯度方向上的两个邻近像素点的灰度值界限内来对图像实行编码。如果该值在该范围内,标记为 1,否则标记为 0。由此获得的 8 bits 二进制代码乘以各位相应的权重,然后累加,其和作为中心像素点 g_c 的 NSBP 特征编码值。NSBP 编码的计算过程如图 2 所示,并且使用式(3)和(4)计算 NSBP 代码。

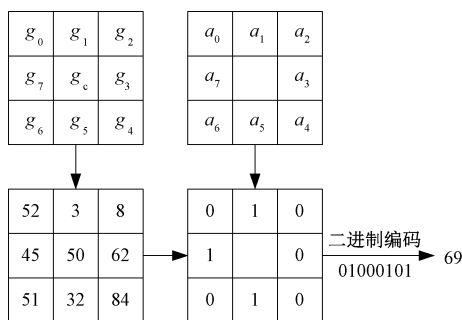


图 2 NSBP 编码计算过程

Fig. 2 The NSBP encoding calculation process

$$\begin{aligned}
 a_0 &= s(g_0, g_1, g_2); a_1 = s(g_7, g_c, g_3) \\
 a_2 &= s(g_6, g_5, g_4); a_3 = s(g_0, g_7, g_6) \\
 a_4 &= s(g_1, g_c, g_5); a_5 = s(g_2, g_3, g_4) \\
 a_6 &= s(g_0, g_c, g_4); a_7 = s(g_2, g_c, g_6)
 \end{aligned}
 \tag{3}$$

$$s(x, y, z) = \begin{cases} 1, & y \in [x, z] \\ 0, & \text{其他} \end{cases}
 \tag{4}$$

式中: a_0, a_1 和 a_2 表示水平方向上的平滑二值模式; a_3, a_4 和 a_5 表示垂直方向上的平滑二值模式; a_6 和 a_7 表示对角线方向上的平滑二值模式; $s(x, y, z)$ 是函数,当 y 在 x 和 z 之间时,该值等于 1,否则为 0。

与整体直方图方法相比,在光照和姿态的变化方面,局部直方图比整体直方图拥有较强的鲁棒性^[15],因此本文采用后者来构建 NSBP 纹理特征的分布。在特征提取步骤中,先将 NSBP 特征图划分成大小一样且相互不重叠子块,接着统计各个子块区域中的直方图分布,最终把它们依照先行后列的次序串联起来以构成块直方图序列。NSBP 模型的降维可以通过使用等效模型来实现^[16]。

1.3 区域特征表示

倘若将一幅面部图像的全部特征都提取,这样并非利于表情识别。如果将所有这些特征都提取不但会增多特征提取和表情分类的时间,而且大批的冗余特征会干预表情类的区分,从而降低表情识别的准确率。所以,合适地选取表情特征区域能够消除少许与表情分类不相干或冗余的特征。针对眉毛、眼睛和嘴巴这 3 个关键区域在表情描绘中的显著作用,本文提取这些区域的 NSBP 纹理特征。

首先把脸部图像规范化成 128×133 区域,其次划分成大小相同的 $56(8 \times 7)$ 个互不重叠小块,这样每个小块的大小为 16×19 。然后,定位眉毛、眼睛和嘴巴区域,其中这 3 个区域大小分别是 $128 \times 19, 96 \times 19, 64 \times 38$ 像素。最后,分别对 3 个区域中每个小块的 NSBP 特征直方图进行计算,同时将这些分块直方图序列串联起来形成区域表情纹理特征。其过程如图 3 所示。

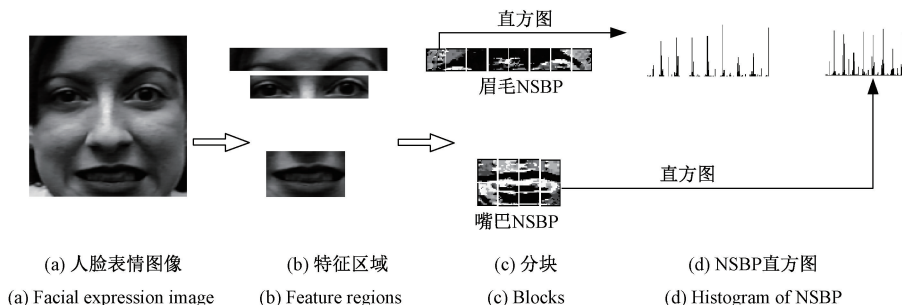


图 3 NSBP 直方图序列

Fig. 3 Histogram sequences of NSBP

2 基于相似性测度的加权证据融合

2.1 基本概率分配

假设 Θ 是某个问题域上互不相容和穷举假设的有

限集合,则 Θ 为辨识框架。如果 $A \subseteq \Theta$,则 $2^\Theta \rightarrow [0, 1]$,且满足式(5)定义的条件。

$$\begin{cases} \sum m(A) = 1 \\ m(\emptyset) = 0 \end{cases}
 \tag{5}$$

式中: m 是 Θ 上的基本概率分配 (basic probability

assignment, BPA) 函数, $m(A)$ 为 A 的支持程度。如果 $m(A) > 0$, 则称 A 为焦元。假设在相同的识别框架 Θ 中有两个相互独立证据, 它们的焦元分别是 B_i 和 $C_j (i = 1, 2, \dots, m; j = 1, 2, \dots, n)$, m 和 n 分别代表两个证据焦元的数目, 这两个证据的 BPA 函数是 m_1 和 m_2 。D-S 组合规则^[17]定义如下:

$$m(A) = \begin{cases} \frac{\sum_{B_i \cap C_j = A} m_1(B_i) m_2(C_j)}{1 - \sum_{B_i \cap C_j = \emptyset} m_1(B_i) m_2(C_j)}, & A \neq \emptyset \\ 0, & A = \emptyset \end{cases} \quad (6)$$

其中, $k = \sum_{B_i \cap C_j = \emptyset} m_1(B_i) m_2(C_j)$ 为冲突系数。表情识别框架 $\Theta = \{c_1, c_2, c_3, c_4, c_5, c_6\}$ 为一个包括 6 种表情的 6 元组。其中, c_1, c_2, c_3, c_4, c_5 和 c_6 分别代表高兴、悲伤、惊讶、厌恶、愤怒和恐惧 6 种不同表情。同时把眉毛、眼睛和嘴巴区域的 NSBP 纹理特征当作实现决策融合的 3 个证据, 它们的基本概率分配函数分别表示成 m_1, m_2 和 m_3 。

在建立这 3 个证据的初始概率赋值时, 首先通过式(7)计算第 i 个区域的第 j 个表情类别的标准表情特征向量。

$$Avg_{ij} = (a_{i1}^j, \dots, a_{ik}^j, \dots, a_{in_i}^j) = \frac{\sum_{l=1}^{n_{c_j}} X_i^{j,l}}{n_{c_j}} \quad (7)$$

$i = 1, 2, 3; j = 1, 2, \dots, 6$

式中: $X_i^{j,l}$ 代表第 i 个区域属于 c_j 的第 l 个训练样本的 NSBP 纹理特征, 而 n_{c_j} 是训练集中属于 c_j 的样本数目。从式(7)可以看出, Avg_{ij} 是第 i 个区域、第 j 个表情类别的聚类中心。

然后通过式(8)计算待测样本和各类标准特征向量之间的卡方距离 $d_{i,j} (i$ 为区域, j 为表情类别)^[18]。

$$d_{i,j} = \sum_{k=1}^{n_i} \frac{(x_{ik} - a_{ik}^j)^2}{x_{ik} + a_{ik}^j} \quad i = 1, 2, 3; j = 1, 2, \dots, 6 \quad (8)$$

其中, $X_i = (x_{i1}, \dots, x_{ik}, \dots, x_{in_i})$ 是测试样本的第 i 个区域的 NSBP 特征。由于距离 $d_{i,j}$ 是 $(0, +\infty)$ 内的任何实数, 则必须通过式(9)进行归一化。

$$d_{i,j} = d_{i,j} / \sum_{k=1}^6 d_{i,k} \quad i = 1, 2, 3; j = 1, 2, \dots, 6 \quad (9)$$

在构建基本概率分配时, 除了每个类别的概率之和应为 1, 同时还必须提高类别概率的区别。若待测样本与某种类型的标准特征向量之间的距离较大时, 待测样本隶属该类别的可能性不大, 指派给该类的概率接近于 0。否则, 分配该类的概率接近于 1。基于上述分析, 本文选取指数函数实现从卡方距离到基本概率分配的映射。该映射关系可由式(10)给出, $m_i(c_j)$ 表示证据 i 对表情类别 c_j 的基本概率分配。

$$m_i(c_j) = \frac{e^{-\lambda d_{i,j}^2}}{\sum_{k=1}^6 e^{-\lambda d_{i,k}^2}} \quad i = 1, 2, 3; j = 1, 2, \dots, 6 \quad (10)$$

式中: λ 为拟合系数, 通过实验取 13。

虽然 D-S 证据理论中的 Dempster 规则可以合并较少冲突的证据, 但是对于冲突较大的证据的合成结果通常是不合理的, 如 Zadeh 悖论^[19]。为了组合冲突较大的证据, 本文基于“证据折扣”的思想, 即根据证据的权重对证据进行折扣以产生新的证据, 然后将其用 D-S 规则进行组合。

2.2 加权证据融合

1) 证据距离

Jousselme 在证据理论中引入了向量空间, 并且在此前提下提出了证据距离的相关概念^[20]。在辨识框架 Θ 中, 假设两个证据的基本概率分配函数是 m_1 与 m_2 , 它们之间的 Jousselme 距离可以表示成:

$$d_{m_1 m_2} = \sqrt{\frac{1}{2} (m_1 - m_2)^T D (m_1 - m_2)} \quad (11)$$

式中: $D[i, j] = |A_i \cap B_j| / |A_i \cup B_j|$ 为 $2^{|\Theta|} \times 2^{|\Theta|}$ 正定系数矩阵; A_i, B_j 分别是 m_1 和 m_2 幂集 2^Θ 中的元素。倘若两个证据之间的距离较大, 那么相互冲突较大, 表示它们之间的相似性较小。两个证据之间的距离可以测量它们之间的冲突程度, 并且解决了冲突系数在一致证据情形中的误判。

由于当证据之间拥有冲突时, 传统 D-S 组合规则会失效, 从而本文提出一种改进的合成方法。首先使用 Jousselme 距离表示证据之间的冲突程度, 并把其转换为证据之间的相似度, 采用相似度决定反映证据重要程度的权重系数。然后根据“折扣率”的构思, 使用权重系数来修正基本概率赋值。最终依然采用 D-S 组合规则融合修正后的 BPA。

2) 加权证据融合

(1) 假设在证据融合过程中, n 个证据源提供证据。证据集为 $E = \{E_1, E_2, \dots, E_n\}$, n 个证据 E_1, E_2, \dots, E_n 对应的基本概率分配函数分别为 m_1, m_2, \dots, m_n 。使用式(11)来统计任意两个证据的 BPA 函数 m_i 与 m_j 的 Jousselme 距离 $d_{m_i m_j}$, 同时表示成冲突矩阵 M_c 。

$$M_c = \begin{bmatrix} 0 & d_{m_1 m_2} & \dots & d_{m_1 m_j} & \dots & d_{m_1 m_n} \\ d_{m_1 m_1} & 0 & \dots & d_{m_1 m_j} & \dots & d_{m_1 m_n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ d_{m_n m_1} & d_{m_n m_2} & \dots & d_{m_n m_j} & \dots & 0 \end{bmatrix}$$

鉴于 $d(m_i, m_j) = d(m_j, m_i)$ 且 $d(m_i, m_i) = 0$, 因此冲突矩阵 M_c 是一个对称矩阵, 同时若任意两个证据之间拥有较小的冲突, 那么它们之间的相似性就越高。证据 m_i 与 m_j 之间的相似度定义如下:

$$S_{ij} = 1 - d_{m_i m_j} \quad i, j = 1, 2, \dots, n \quad (12)$$

(2) 利用式 (12) 将冲突矩阵 M_c 转化成相似矩阵

M_S 。

$$M_S = \begin{bmatrix} 1 & S_{12} & \dots & S_{1j} & \dots & S_{1n} \\ S_{i1} & S_{i2} & \dots & S_{ij} & \dots & S_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ S_{n1} & S_{n2} & \dots & S_{nj} & \dots & 1 \end{bmatrix}$$

(3) 根据相似矩阵 M_S , 证据 m_i 被其他证据支持的程度 $Sup(m_i)$ 由式 (13) 给出。

$$Sup(m_i) = \sum_{j=1, j \neq i}^n S_{ij} \quad i = 1, 2, \dots, n \quad (13)$$

$Sup(m_i)$ 为计算相似矩阵 M_S 中各行除其本身相似度外的其他全部元素的总和, 同时它能够反映出 m_i 被其他证据所支持的程度, 类似于度量函数。若一个证据和其他证据的相似度较高, 就认为它们互相支持对方; 反之则认为它们相互支持的程度也低。

(4) 对融合系统中每个证据的 $Sup(m_i)$ 实行排序。将具有最大 $Sup(m_i)$ 的证据当作关键证据, 而其余证据均是非关键证据。每个证据相对于关键证据的证据权重可通过式 (14) 获得。

$$w_i = \frac{Sup(m_i)}{\max_{1 \leq j \leq n} \{Sup(m_j)\}} \quad i = 1, 2, 3 \quad (14)$$

根据“折扣率”的思想, 运用上述计算的证据权重来转化各证据的初始概率赋值。令 m'_i 表示 m_i 调整后的基本概率分配, 其转换方法如下:

$$m'_i(A) = w_i m_i(A), \quad \forall A \in 2^\Theta, A \neq \Theta \quad (15)$$

$$m'_i(\Theta) = (1 - w_i) + w_i m_i(\Theta) \quad (16)$$

从式 (15) 可以看出, 权重较小的证据, 也就是相互支持度不高的证据元素 A 给出的确定性信息会降低, 而式 (16) 会增加权重较小的证据的不确定性元素 Θ 提供的不确定性信息。从而能够减少互相支持程度不高的证据对总体合成结果的影响。

(5) 最终仍然采用 D-S 组合规则合成变换后的 BPA 函数 m'_i , 其定义如下:

$$m(A) = \begin{cases} \frac{\sum_{\cap A_i = A} \prod_{i=1}^n m'_i(A_i)}{1 - \sum_{\cap A_i = \emptyset} \prod_{i=1}^n m'_i(A_i)}, & A \neq \emptyset \\ 0, & A = \emptyset \end{cases} \quad (17)$$

改进后的组合公式与传统的 D-S 组合公式相同, 从而保持了 D-S 组合规则的优点。 $\sum_{\cap A_i = \emptyset} \prod_{i=1}^n m'_i(A_i)$ 是重估的系数 k 。各个证据的权重系数 w_i 是由 Jusselme 距离决定, 同时解决了冲突系数量化冲突存在的情形。此外, 证据权重系数的大小能够体现证据之间相互支持程度的高低。依据权重重新调整基本概率赋值, 并减少证据相互之间的冲突程度以符合 D-S 融合规则的条件。

至此, 本文提出的基于区域 NSBP 特征的加权证据

融合的面部表情识别框架如图 4 所示。

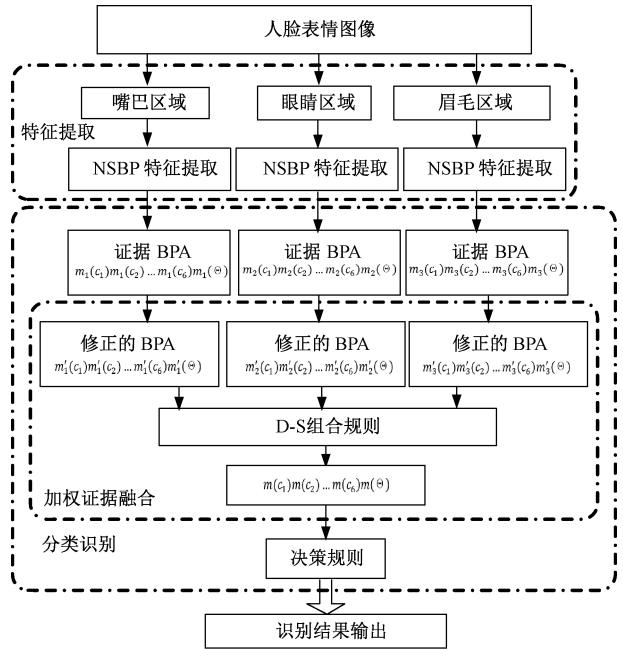


图 4 表情识别框架

Fig. 4 The framework of expression recognition

3 实验结果与分析

为了证明本文算法的有效性, 利用 MATLAB 2010 在 Windows 10 环境下构建系统, 计算机微处理器配置为 Intel(R) Core(TM) i3-4150 CPU@ 3.5GHZ、内存 4 GB。本文实验采用 Cohn-Kanade(CK) 人脸表情数据库。为了更好地减少图像中人脸位置, 大小和旋转等因素对表情识别的影响, 对图像实行一系列预处理工作, 如人脸分割、几何归一化和光照归一化。预处理后, 表情图像的统一尺寸为 128×133, 如图 5 所示。



图 5 CK 表情库的预处理

Fig. 5 Preprocessing on CK expression database

3.1 本文算法识别率

CK 库包含 100 名大学生的面部表情图像。每个人都有一组图像序列, 这些图像序列是从中性到某些面部动作单元编码的面部显示。本文实验从 CK 数据库中选

择了来自 21 个对象的 6 种表情(高兴、悲伤、惊讶、厌恶、愤怒和恐惧)的 200 幅人脸图像。每次任意选择 100 幅图像当作训练集合,其余图像作为待测集合。

对于随机生成的表情训练样本,提取各个训练图像的 3 个关键区域的 NSBP 纹理特征后,将各个区域中相同类型的表情特征累加并计算平均值,以获得各个区域不同类的标准表情特征向量。在测试过程中,提取各个待测图像的 3 个关键区域的 NSBP 特征之后,分别计算 3 个区域特征与标准表情特征向量之间的卡方距离,以构造 3 个区域证据的基本概率分配。最终根据权重对基本概率分配函数实行调整,同时完成证据融合和最后决策。4 次交叉验证的平均识别率是 95.25%,其结果如表 1 所示。

表 1 基于 CK 数据库的实验结果

Table 1 Experimental 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	第 4 次	
高兴	14	13	14	14	14	98.21
悲伤	15	13	14	14	14	91.67
惊讶	15	15	15	15	14	98.33
厌恶	19	18	19	18	19	97.37
愤怒	18	17	17	16	17	93.06
恐惧	19	18	17	18	18	93.42
总计	100	94	96	95	96	95.25

该方法在 CK 数据库上的混淆矩阵如表 2 所示。高兴、惊讶和厌恶的表情很容易识别,而其余的表情则不然。同时,大多数混淆发生在愤怒和厌恶(4.16%),悲伤和愤怒(3.33%)以及恐惧和惊讶(2.63%)的表情之间。分析原因:愤怒和厌恶表情在嘴巴周围的纹理改变较类似,悲伤和愤怒表情纹理变化都较弱,恐惧和惊讶两种表情在眼睛附近的纹理较为相似。

表 2 本文方法在 CK 数据库上的混淆矩阵

Table 2 Confusion matrix for the proposed method on CK database

	高兴	悲伤	惊讶	厌恶	愤怒	恐惧
高兴	98.21	0.00	0.00	0.00	0.00	1.79
悲伤	0.00	91.67	0.00	0.00	3.33	5.00
惊讶	0.00	0.00	98.33	0.00	1.67	0.00
厌恶	1.31	0.00	0.00	97.37	1.32	0.00
愤怒	0.00	1.39	0.00	4.16	93.06	1.39
恐惧	1.32	0.00	2.63	1.32	1.31	93.42

3.2 不同表情识别方法比较

为了说明 NSBP 特征比 LBP 特征可以更好地表示局部细节信息,并检验 WEF 优于 D-S,本文对于不同的纹

理特征算子和分类器进行了实验,它们之间的比较结果如表 3 所示。表 3 表明 NSBP 特征和 WEF 的结合可以获得更高的识别率。

表 3 不同特征和不同分类器的平均识别率比较

Table 3 Comparison of the average recognition rates of different features and different classifiers

特征	分类器	平均识别率/%
LBP	D-S	88.53
NSBP	D-S	92.07
LBP	WEF	92.28
NSBP	WEF	95.25

为了更好地说明本文方法在表情识别率和识别时间方面的优点,对于一样的数据集合,把该方法和其他表情识别方法做比较。同时这些方法的关联参数都依照原文默认配置,其比较结果如表 4 所示。

表 4 不同表情识别方法的性能比较

Table 4 Performance comparison of different expression recognition methods

表情识别方法	特征提取时间+ 识别时间/ms	平均识别率/%
HOG+SVM ^[5]	693+1 542	93.14
LBP+kNN+D-S ^[12]	678+1 623	93.25
LBP+SVM ^[21]	656+1 351	92.46
Gabor+FLD+kNN ^[10]	38 490+651	93.78
Gabor+Fisher+SVM ^[22]	35 600+1 356	92.92
LBP+Gabor+SVM ^[23]	40 120+2 434	96.73
NSBP+WEF	721+765	95.25

表 4 为 LBP 和 HOG 算法的特征提取时间基本上是相同的数量级。就特征提取时间而言,鉴于 NSBP 特征不复杂的计算过程,该方法略微高于传统的 LBP 算子,然而与多尺度多方向 Gabor 特征相比,具有更好的优势。就识别时间而言,由于 SVM 在训练步骤中存在计算需要大量时间的情况,而 WEF 不需要繁琐的训练步骤即可获得模型参数,因而拥有较短的识别时间。就分类性能而言,因为提出了具有较强局部细节表示的 NSBP 算子,因此基于 NSBP 算子的分类性能要优于单一特征的 HOG、LBP 和 Gabor 表情识别方法。尽管表情识别率不高于 LBP + Gabor 混合特征方式,但本文方法在特征提取和表情识别时间方面具有更好的优势。经过衡量整体分类时间和分类性能,经过提取关键区域的 NSBP 纹理特征同时利用 WEF 来融合区域特征,本文方法不但可以获得较高的识别率,而且拥有较短的识别时间。所以,所提出的方法可以有效地符合诸如人机交互的实时情景。

4 结 论

提出了一种基于区域 NSBP 特征及 WEF 的人脸表情识别方法。和类似方法在识别性能上作比较,不但能够确保较高的表情识别率,而且还可以减少识别时间。提出的方法的具体性能如下。

1) 针对 LBP 仅计算中心点和邻域点之间的大小关系的局限,提出的 NSBP 纹理算子方法是通过判别水平、垂直及对角线这 3 个方向的“中心”像素点的灰度值是不是在每个梯度方向上两个邻域的灰度值的范畴之间对图像实行编码。因此可以提取更有鉴别力的纹理特征。

2) 基于 WEF 策略,能够体现证据重要程度的权重系数取决于证据之间的互相支持程度,同时采用权重系数对基本概率分布进行调整,通过权重反映不同证据的重要性。它不但可以降低权重较小的证据对决策判断的影响,而且还能减少修正后证据之间的相互冲突,从而满足使用 D-S 组合规则的条件。

本文所提出的方法是针对单个静态图像和正面面部的表情识别,然而人脸表情识别的过程通常不是静态的。未来的研究应该确定动态信息的有效使用,以提高姿势变化大的图像和视频序列中的表情识别率。

参考文献

- [1] 徐桂芝, 赵阳, 郭苗苗, 等. 基于深度分离卷积的情绪识别机器人即时交互研究 [J]. 仪器仪表学报, 2019, 40(1):161-168.
X G Z, ZHAO Y, GUO M M, et al. Research on real-time interaction for the emotion recognition robot based on depth-wise separable convolution [J]. Chinese Journal of Scientific Instrument, 2019, 40(1): 161-168.
- [2] MUNIR A, HUSSAIN A, ALIKHAN S, et al. Illumination invariant facial expression recognition using selected merged binary patterns for real world images [J]. Optik, 2018, 158:1016-1025.
- [3] FADI D, ABDELMALIK M, BOGDAN R. Facial expression recognition using tracked facial actions: Classifier performance analysis [J]. Engineering Applications of Artificial Intelligence, 2013, 26(1): 467-477.
- [4] ZHOU H L, LAM K M, HE X J. Shape-appearance-correlated active appearance model [J]. Pattern Recognition, 2016, 56(C):88-99.
- [5] CHEN J K, CHEN Z H, CHI Z R, et al. Facial expression recognition based on facial components detection and HOG features [C]. International Workshops on Electrical and Computer Engineering, 2014:64-69.
- [6] HEGDE G, SEETHA M, HEGDE N. Facial expression recognition using entire Gabor filter matching score level fusion approach based on subspace methods [C]. International Conference on Mining Intelligence and Knowledge Exploration, 2015:47-57.
- [7] 胡敏, 朱弘, 王晓华, 等. 基于梯度 Gabor 直方图特征的表情识别方法 [J]. 计算机辅助设计与图形学学报, 2013, 25(12): 1856-1861.
HU M, ZHU H, WANG X H. Expression recognition method based on gradient histogram features [J]. Journal of Computer-Aided Design & Computer Graphics, 2013, 25(12): 1856-1861.
- [8] ZHAO X M, ZHANG S Q. Facial expression recognition based on local binary patterns and kernel discriminant isomap [J]. Sensors, 2011, 11(10):9573-9588.
- [9] 吴昊, 胡敏, 高永, 等. 融合 DCLBP 和 HOAG 特征的人脸表情识别方法 [J]. 电子测量与仪器学报, 2020, 34(2):73-79.
WU H, HU M, GAO Y, et al. Facial expression recognition based on DCLBP and HOAG [J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(2):73-79.
- [10] GU W F, 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.
- [11] 王晓华, 金超, 任福继, 等. 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.
WANG X H, JIN C, REN F J, et al. 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.
- [12] SHOYAIB M, ABDULLAH-AI-WADUD M, ISHRAQUE S M Z, et al. Facial expression classification based on Dempster-Shafer theory of evidence [C]. International Conference on Belief Functions, 2012:213-220.
- [13] 杨晓萍, 廉伟健, 李孟杰, 等. 基于 Pignistic 概率函数和相关系数的证据组合方法 [J]. 通信学报, 2019, 40(3):73-79.
YANG X P, LIAN W J, LI M J et al. Evidence combination method based on Pignistic function transformation and correlation coefficient [J]. Journal on Communications, 2019, 40(3):73-79.
- [14] 汤兰兰, 盖绍彦, 达飞鹏, 等. 基于网格纵横局部二值模式的三维人脸识别 [J]. 仪器仪表学报, 2016, 37(6):1413-1420.

- TANG L L, GAI S Y, DA F P, et al. 3D Face recognition method based on the local binary pattern from vertical and horizontal on the mesh [J]. Chinese Journal of Scientific Instrument, 2016, 37(6):1413-1420.
- [15] 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 CH. Facial expression recognition method based on Gabor multi-orientation features fusion and block histogram [J]. Acta Automatica Sinica, 2011, 37(12):1455-1463.
- [16] 田思洋, 徐科, 郭会昭. 局部二值模式在连铸坯表面缺陷识别中的应用[J]. 工程科学学报, 2016, 38(12):1728-1733.
TIAN S Y, XU K, GUO H ZH. Application of local binary patterns to surface defect recognition of continuous casting slabs [J]. Chinese Journal of Engineering, 2016, 38(12):1728-1733.
- [17] XIAO F Y. Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy [J]. Information Fusion, 2019(46):23-32.
- [18] TAN X Y, TRIGGS B. Enhanced local texture feature sets for face recognition under difficult lighting conditions [J]. IEEE Transactions on Image Processing, 2010, 19(6):1635-1650.
- [19] 薛大为, 王永, 高康凯. 基于规范分解的证据合成悖论分析[J]. 北京邮电大学学报, 2019, 42(1):28-34.
XUE D W, WANG Y, GAO K K. Analysis of evidence combination paradox based on canonical decomposition [J]. Journal of Beijing University of Posts and Telecommunication, 2019, 42(1):28-34.
- [20] 周永庆, 韩德强, 杨艺. 证据距离的选取对冲突证据组合影响的研究[J]. 西安交通大学学报, 2018, 52(6):1-8.
ZHOU Y Q, HAN D Q, YANG Y. A research on the influence of evidence distance selection on combination of conflict evidences [J]. Journal of Xi'an Jiaotong University, 2018, 52(6):1-8.
- [21] ZHONG L, LIU Q S, YANG P, et al. Learning active facial patches for expression analysis [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2012:2562-2569.
- [22] LIU H B, ZHANG G B, Huang Y M, et al. Multiple features extraction and coordination using Gabor wavelet transformation and fisherfaces with application to facial expression recognition [C]. Chinese Conference on Pattern Recognition, 2010:1-5.
- [23] ZAVASCHI T H H, BRITTO A S, OLIVEIRA L E S, et al. Fusion of feature sets and classifiers for facial expression recognition [J]. Expert Systems with Application, 2013, 40(2):646-655.

作者简介



刘娟, 2009 年于合肥工业大学获得硕士学位, 现为安庆师范大学副教授, 主要研究方向为情感计算、机器视觉和自然语言处理。

E-mail: juanliu3039@163.com

Liu Juan received her M. Sc. degree in 2009 from Hefei University of Technology. Now she is an associate professor at Anqing Normal University. Her main research interests include affective computing, machine vision and natural language processing.



胡敏, 2004 年于合肥工业大学获得博士学位, 现为合肥工业大学教授、博士生导师, 主要研究方向为计算机视觉、情感计算等。

E-mail: uhnim@163.com

Hu Min received her Ph. D. degree from Hefei University of Technology in 2004. Now she is a professor and Ph. D. supervisor at Hefei University of Technology. Her main research interests include computer vision, affective computing and so on.