

基于多尺度细节增强的面部表情识别方法

谭小慧^{①③} 李昭伟^{①④} 樊亚春^{*②}

^①(首都师范大学信息工程学院 北京 100048)

^②(北京师范大学虚拟现实应用教育部工程研究中心 北京 100875)

^③(首都师范大学电子系统可靠性与数理交叉学科国家国际科技合作示范型基地 北京 100048)

^④(首都师范大学北京成像理论与技术高精尖创新中心 北京 100048)

摘 要: 人类面部表情是其心理情绪变化的最直观刻画, 不同人的面部表情具有很大差异, 现有表情识别方法均利用面部统计特征区分不同表情, 其缺乏对于人脸细节信息的深度挖掘。根据心理学家对面部行为编码的定义可以看出, 人脸的局部细节信息决定了其表情意义。因此该文提出一种基于多尺度细节增强的面部表情识别方法, 针对面部表情受图像细节影响较大的特点, 提出利用高斯金字塔提取图像细节信息, 并对图像进行细节增强, 从而强化人脸表情信息。针对面部表情的局部性特点, 提出利用层次结构的局部梯度特征计算方法, 描述面部特征点局部形状特征。最后, 使用支持向量机(SVM)对面部表情进行分类。该文在CK+表情数据库中的实验结果表明, 该方法不仅验证了图像细节对面部表情识别过程的重要作用, 而且在小规模训练数据下也能够得到非常好的识别结果, 表情平均识别率达到98.19%。

关键词: 表情识别; 图像金字塔; 高斯差分; 细节增强; 支持向量机

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2019)11-2752-08

DOI: 10.11999/JEIT181088

Facial Expression Recognition Method Based on Multi-scale Detail Enhancement

TAN Xiaohui^{①③} LI Zhaowei^{①④} FAN Yachun^{*②}

^①(College of Information Engineering, Capital Normal University, Beijing 100048, China)

^②(College of Information Science and Technology, Beijing Normal University, Beijing 100875, China)

^③(Beijing Key Laboratory of Electronic System Reliability and Prognostics,
Capital Normal University, Beijing 100048, China)

^④(Beijing Advanced Innovation Center for Imaging Technology,
Capital Normal University, Beijing 100048, China)

Abstract: Facial expression is the most intuitive description of changes in psychological emotions, and different people have great differences in facial expressions. The existing facial expression recognition methods use facial statistical features to distinguish among different expressions, but these methods are short of deep exploration for facial detail information. According to the definition of facial behavior coding by psychologists, it can be seen that the local detail information of the face determines the meaning of facial expression. Therefore, a facial expression recognition method based on multi-scale detail enhancement is proposed, because facial expression is much more affected by the image details than other information, the method proposed in this paper extracts the image detail information with the Gaussian pyramid firstly, thus the image is enhanced in detail to enrich the facial expression information. Secondly, for the local characteristics of facial expressions, a local gradient feature calculation method is proposed based on hierarchical structure to describe the local shape features of facial

收稿日期: 2018-11-26; 改回日期: 2019-02-27; 网络出版: 2019-05-20

*通信作者: 樊亚春 fanyachun@hotmail.com

基金项目: 国家重点研发计划项目(2017YFB1002804), 国家自然科学基金项目(61602324), 浙江大学CAD&CG国家重点实验室开放课题(A1914)

Foundation Items: The National Key R&D Program of China(2017YFB1002804), The National Natural Science Foundation of China(61602324), The Open Project of State Key Lab of CAD & CG, Zhejiang University(A1914)

feature points. Finally, facial expressions are classified using a Support Vector Machine (SVM). The experimental results in the CK+ expression database show that the method not only proves the important role of image detail in facial expression recognition, but also obtains very good recognition results under small-scale training data. The average recognition rate of expressions reaches 98.19%.

Key words: Expression recognition; Image pyramid; Gauss difference; Detail enhancement; Support Vector Machine(SVM)

1 引言

面部表情识别是计算机图形图像学和人机交互领域的一个重要研究问题, 其应用领域十分广泛, 可应用于司法、安全、医疗、通信、汽车制造等不同领域。例如: 面部表情识别可应用于刑事犯罪侦查中, 利用犯罪嫌疑人的脸部表情推测其心理变化; 可应用于虚拟购物, 根据用户脸部表情识别推断其喜好, 进而推荐符合用户需求的各项商品服务; 可应用于在线学习, 根据学习者脸部表情变化推断其对所学习内容的兴趣程度, 引导用户进行有效性学习; 可应用于智能虚拟人交互, 通过用户情感识别, 调整虚拟人的表情, 达到自然人机交互的目的。

人脸表情中含有丰富的信息, Mehrabian^[1]提出人类情感信息有55%是通过人脸传递的。如果使计算机能够读取面部表情信息, 那么在人机交互过程中会有更好的体验。Ekman对人脸表情运动深入分析并定义面部行为编码系统(Facial Action Coding System, FACS), 该编码系统针对面部肌肉块的运动规律, 定义了不同种类的人脸运动单元描述面部不同表情, 人脸运动单元表达了面部局部细节信息的变化。不同运动单元的组合形成了人脸的不同类型的表情, Ekman将人脸表情区分为6种基本表情: 愤怒、厌恶、惊讶、高兴、悲伤、恐惧。

目前, 在Ekman的研究基础上, 面部表情识别已经在情感计算方面取得了很大进步。面部表情识别一般可以分为3个步骤: 图像预处理、表情特征表达和表情识别。表情特征表达是面部表情识别过程中最为关键的一步, 直接影响表情识别的效果。面部特征主要可以分为: 纹理特征(texture feature)和几何特征(geometric feature)。局部二值模式(Local Binary Pattern, LBP)以其简单快速的优点被学者们引入到面部特征提取中: Guo等人^[2]提出了完全局部二值模式(Completed LBP, CLBP)的纹理分类方法, 该方法深度分析了LBP方法, 并提出中心描述子、符号描述子和大小描述子等3种描述子, 然而, 未能解决CLBP对高斯噪声敏感的问题。王伟等人^[3]提出LBP金字塔特征, 提取各层金字塔图像的LBP特征, 分块统计各层的LBP特征, 然后级联成图像特征。Gabor特征具有方向不

变性, 因此在传统面部特征提取方法中占有比较重要的地位: 钟思志^[4]提出的基于Gabor特征的特征值提取法, 利用局部分类器对局部特征进行分类, 加强局部特征提取, 然而该方法对图片噪声和强光的鲁棒性较低。杨凡等人^[5]采用定义Gabor参数矩阵与改进Adaboost分类器相结合的方法, 构建最强分类器, 从而降低Gabor特征的维度和冗余。基于方向梯度直方图(Histogram of Oriented Gradient, HOG)的梯度统计方法也是一大研究点: 童莹^[6]针对HOG特征忽略局部特征之间信息关系的问题, 将图像划分为一系列不同的子区域, 统计子区域的HOG特征形成整幅图像的HOG特征。Turan等人^[7]将面部划分为嘴和眼睛等窗口, 采用局部相位量化(Local Phase Quantization, LPQ)和金字塔式梯度方向直方图(Pyramid of Histogram of Oriented Gradients, PHOG)描述局部特征信息, 使用典型相关分析(Canonical Correlation Analysis, CCA)对特征表示, 有效地去除了冗余信息; Saeed等人^[8]通过对面部主要位置定位8个关键点, 计算获得6个归一化几何特征向量作为几何特征, 虽然取得了不错的效果, 因为关键点位置较少, 忽略了微小表情信息。Chen等人^[9]使用正交方向梯度直方图(Histogram of Oriented Gradients from Three Orthogonal Planes, HOG-TOP)提取纹理特征, 在面部运动过程中, 根据面部特征点变化, 提取面部运动的几何信息, 在CK+库上验证了该方法的有效性。任福继等人^[10]提取图像的局部二值模式—3维正交平面(Local Binary Patterns on Three Orthogonal Planes, LBP-TOP)和梯度方向直方图—3维正交平面(HOG-TOP)级联, 同时提取图像的BVP生理信号, 共同作为表情特征的描述, 使用BP分类训练模型, 模糊积分决策的方式识别表情类别, 证明了方法的有效性, 然而存在训练过程复杂的问题。

除了传统方法外, 卷积神经网络(Convolutional Neural Networks, CNN)也是一种比较流行的方法: 在未有大数据量的数据库背景下, Lopes等人^[11]为了解决卷积神经网络需要大数据量提高精度的问题, 提出将CNN和图片预处理相组合的方法, 在

CK+等库中验证了该方法的有效性。Zhang等人^[12]受深度身份网络(Deep IDentity Network, DeepID-Net)启发,提出身份启发网络(Identity-Inspired CNN, I²CNN),提取整个或者局部面部的微小区域作为I²CNN的输入,使用SVM作为分类器,取得很好的表情识别效果,然而,该方法未能解决CNN参数耗时的问题,算法受图像人物的光照、姿势和种族等因素影响。虽然,CNN在情感计算上取得了不错的研究进展,但是CNN存在训练过程需要大量的训练样本、消耗时间长、实现复杂等缺点。本文采用的高斯平滑对噪声处理也起到重要作用,梯度具有整体光照不变性,采用4个方向统计梯度信息的方式,降低了计算量,使用图像金字塔,丰富了图像信息。

计算机识别面部表情,需要表情数据样本作为训练集,训练出稳定的、可靠的表情识别模型。目前,自然条件下的面部表情数据库较难获取,这使得现有的面部表情识别模型存在着不确定性。Lucey等人^[13]于2010年发布了 The extended Cohn-Kanade dataset (CK+),该库是由在18到30岁之间的123个人的表情序列组成,受试者中65%是女性,囊括了黑色、白色和黄色3种皮肤的人种,每组序列是从中性表情图像到表情顶点图像,表情标签是数据库创造者经过严格验证。JAFPE (Japanese Female Facial Expression database)是10位日本女性的正面表情的数据库,在6种基本表情的基础上加入了中性表情,共213张图像,因此,JAFPE数据库有性别、人种等方法缺陷。本文方法在肤色、性别、年龄等数据较为丰富的CK+表情数据库进行测试。

本文提出多尺度细节增强面部识别的方法。为了减小图像大小、旋转和背景影响,使用Haar特征(Haar-like特征)检测面部和监督下降法(Supervised Descent Method, SDM)对面部使用特征点标记,特征点不随图像的大小、旋转等变换而改变,同时为特征提取排除了背景影响。对图像进行采样形成图像金字塔,获取不同尺寸的表情缩放图像,降低了图像大小对信息提取的不利影响,丰富了特征点周围的特征信息。高斯模糊具有降低细节层次的作用,对输入图像进行多层次的高斯模糊,不同的高斯模糊图像做差,提取图像的细节信息,然后和高斯模糊的输入图像叠加,实现对图像边缘细节信息增强,提高图像数据表达的能力。根据梯度对光照的整体不变性,统计和计算细节增强图像多层次局部梯度直方图,形成表情特征描述子。最后,根据训练集的表情特征描述子训练出SVM表情分类模型,通过SVM表情分类模型对输入的待测特征描述子进行表情预测。

2 多尺度细节增强的人脸表情识别

考虑到人脸细节信息对于表情识别的重要性,本文提出基于多尺度细节增强的人脸表情识别方法,本文方法的流程如图1所示,整个方法分为3个主要部分:首先,对待测样本进行图像预处理。利用SDM人脸对齐方法对检测到的人脸进行特征点检测,获取人脸特征区域的51个特征点。其次,对待测样本提取特征描述子。将输入图像转化为灰度图像,使用采样方法构造人脸图像金字塔,采用高斯模糊对每层金字塔图像进行处理并差分得到高斯模糊图像差分(DoG)图像,获得人脸细节图像,与该层图像进行叠加,形成细节增强图像。统计增强图像的局部多层次梯度直方图形成特征描述子。最后,训练SVM分类模型以及对特征描述子分类预测,识别待测样本的表情。

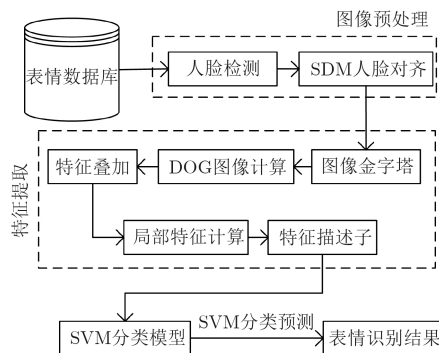


图1 多尺度细节增强人脸表情识别流程

2.1 特征点标记

在图像预处理过程中,本文用SDM算法对人脸建立坐标系,利用监督下降法解决非线性最小化问题,解决了方程不可微、方程计算量太大、Hesse矩阵太大、Hesse矩阵不是正定矩阵等问题。

为了准确检测不同面部特征点的准确位置,需要利用样本人脸中手工标注的特征点对平均人脸特征点的位置进行最优化回归计算。定义优化函数为

$$f(x) = \|H(A(x_0 + \Delta x)) - H(A(x_0))\|^2 \quad (1)$$

其中, $A(x_0)$ 为平均人脸的特征点,平均人脸特征点是对样本人脸中的特征点位置进行平均计算得到。 $A(x_0 + \Delta x)$ 为当前人脸特征点。 H 函数用来计算特征点局部区域特征,本文取HOG特征。

面部表情信息集中在人的嘴巴、鼻子、眼睛、眉毛周围区域传递,人脸的边缘、脸颊、耳朵等区域基本不反映人脸的表情信息,为了保证采集到的图像信息具有有效性,减少数据维数,提高效率,本文围绕表情集中区域标记了51个特征点。通过训练SDM模型,实现人脸特征点自动提取。特征点

分布如图2所示,左眉和右眉各有5个特征点定位,选取9个特征点覆盖鼻子的主要部分,左眼和右眼各设有6个特征点,嘴唇部位分布20个特征点,完成面部部位的定位。



图2 特征点分布图

2.2 图像细节增强

为了使算法对人脸缩放具有不变性,减少人脸缩放变化对于形状特征的影响,本文利用建立图像高斯尺度空间的方法,形成图像金字塔结构。图像金字塔是一系列以金字塔形状排列的、分辨率逐步降低、来源于同一张原始图的图像集合。图像金字塔是多尺度计算的基础,以原始图像的灰度图像为基础,选取该灰度图像邻近的上采样图像和下采样图像,与灰度图像形成3层的图像金字塔,如图3(a)所示的图像金字塔结构,下采样图样是原始图像的1/4,上采样图像是原始图像的4倍。

人脸细节信息对于面部表情的识别起着不可替代的作用。人脸表情的细微变化通常体现在图像细节上,是局部细节累积的结果。高斯拉普拉斯算子(Laplace of Gaussian function, LoG)能够得到图像局部细节,但高斯拉普拉斯算子要进行离散数字图像的2阶差分或者连续情况下的2阶偏导数,增加了计算量。而不同尺度的高斯模糊图像差分(DoG)

能够近似的表征LoG算子,因此本文提出基于DoG的图像局部细节计算方法。

利用高斯平滑具有降低图像细节层次的作用,不同尺度的高斯平滑对图像的平滑效果不同,对图像金字塔的每一层图像做3次不同的高斯平滑,形成3层高斯平滑图像

$$L_i(x, y) = G_i * I(x, y), i = 1, 2, 3 \quad (2)$$

其中, $G_i = \exp(-x^2 - y^2/2\sigma_i)$, *表示卷积操作。

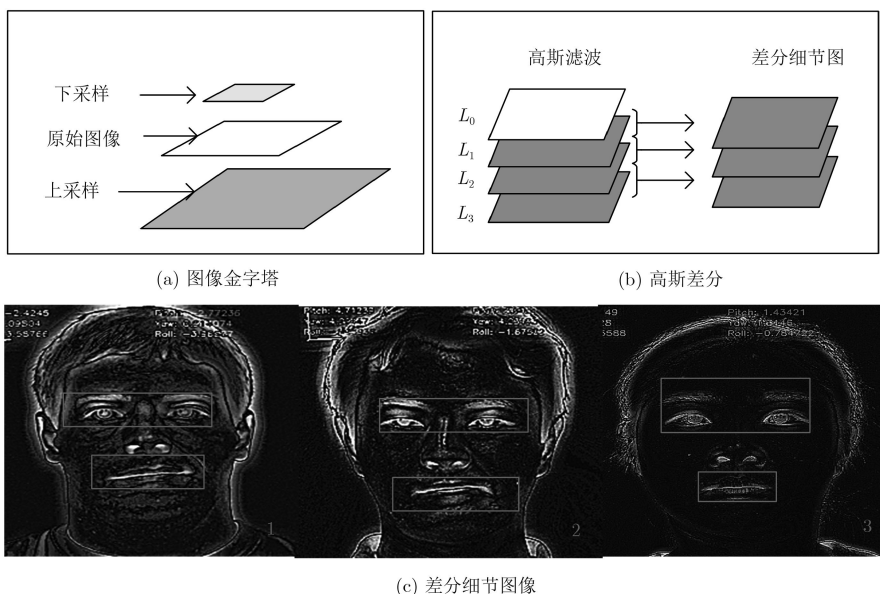
如图3(b)所示,取 L_0 为输入高斯平滑的金字塔图像,与该层图像处理后得到的3层平滑图像形成4层的图像,相邻滤波图像相减得到DoG图像

$$D_i = L_{i-1} - L_i, i = 1, 2, 3 \quad (3)$$

每一层的金字塔图像产生3层DoG图像,如图3(c)中差分细节图像所示的3张图像,是3张测试图像经过计算后,得到每张图像在式(3)中 $i = 2$ 的DoG图像 D_2 ,从图3(c)中可以看出, D_2 图像提取了图像的边缘信息,尤其是面部表情集中表达区域(眼睛和嘴)的细节信息,在整个面部中, D_2 图像的面部器官具有明显的皮肤褶皱信息。

高斯差分得到的图像是在不同尺度下的图像局部细节描述,图像的细节描述能够提高人脸图像对于细节的表达能力。将每层金字塔图像与该层图像的3层DoG图像进行像素叠加,图像中的像素点的最大值是255,为了防止溢出,取图像叠加的平均值,形成每层图像金字塔图像的细节叠加图像,如图4所示, $j = 0, 1, 2$ 。叠加公式为

$$L = \frac{1}{4} \left(L_0 + \sum_{i=1}^3 D_i \right) \quad (4)$$



(c) 差分细节图像
图3 高斯图像差分

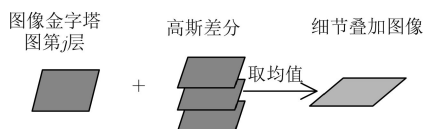


图4 细节特征叠加图像

经过细节增强后的人脸图像,其细节信息更丰富,提取出的局部特征更具有代表性,如图5所示,细节增强前的图5(a)和细节增强后的图5(b),图5(b)覆盖区域为细节增强区域,对比图2,图5中覆盖区域集中在特征点分布区域,通过细节叠加,提高细节表达能力。



(a) 细节增强前



(b) 细节增强后

图5 细节增强前后对比

2.3 层次特征提取与表情识别

根据梯度具有整体光照不变性,通过计算以每一个特征点为中心的像素域内的梯度,形成梯度直方图,作为情感分类的特征描述子。SDM方法定位51个特征点的梯度直方图组合形成人脸图像的表情特征描述子。

本文对于每个特征点选取以其为中心的 16×16 的局部区域,并将此 16×16 区域分为大小不同的4层结构,每层结构为不同数量的矩形块组成。4层结构分别是16个 4×4 的块,9个 8×8 的块,4个 12×12 的块以及1个 16×16 的块。如图6所示。

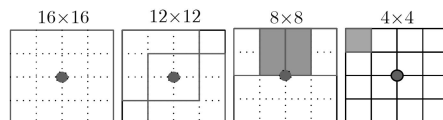


图6 像素区域

图7所示,分布在面部的嘴唇、鼻子、眼睛和眉毛的51个特征点,以每个特征点为中心的像素块基本上覆盖面部表情表达的主要区域,能够比较完整地提取这些特征部位的特征信息。

对每一个块内所有像素计算梯度方向及幅值:块内的像素梯度按照 $0^\circ, 90^\circ, 180^\circ, 270^\circ$ 4个方向进行累加统计,形成4个值的直方图,特征点区域内的所有梯度直方图形成特征点的描述子 f_i 。所有的特征点的描述子共同形成人脸表情的特征描述子 F

$$F = \prod_{i=1}^{51} f_i \quad (5)$$

本文使用SVM方法,根据输入的训练集样本的特征描述子 F 和表情类别,训练出表情分类模型。使用SVM分类模型将待测特征描述子 F 映射到训练的模型空间中分类,从而识别表情。

3 实验结果与分析

3.1 实验数据集

本文使用CK+表情数据库提供的表情图像进行实验,CK+数据库是由卡耐基梅隆大学的Lucey等人^[13]于2010年发布,其中包含123个人的表情序列,每组表情序列是从表情的中性到表情表达的最大程度。在2018年,Yao等人^[14]采用初始、顶峰和结束3种状态描述每一种表情的图像序列,Yao等人使用图像序列的第1帧作为中性表情。本文根据Lucey等人^[13]对CK+的介绍,并借鉴文献^[14]的方法,筛选不同人种、不同类别的表情序列中的中性表情,测试了愤怒、厌恶、恐惧、高兴、悲伤、惊讶和中性表情在内的7种表情,并选取每个分类序列的后

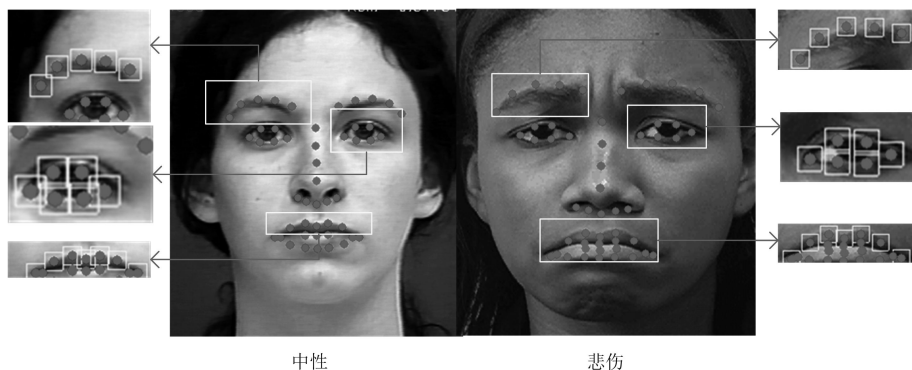


图7 特征点及区域图

5张图像作为相应表情分类数据，并随机抽取50%的图像作为样本训练数据，其余50%作为测试数据。如图8所示，数据集中包含的7种表情示例图。



图8 CK+表情库7种表情示例图像

3.2 实验结果及分析

为了排除测试集和训练集的偶然性，对数据集进行5次随机抽样，并对5次实验结果取平均值。

为了说明高斯模糊参数在图像细节叠加中的影响，使系统识别效果达到最优，本文实验探究高斯模糊参数对系统识别效率的影响。

高斯模糊半径取区间[1, 25]中的奇数值，根据经验，三尺度高斯模糊的标准差 σ 参数取 $\sigma_0(21, 21, 21)$ ，表1为整体识别率最高的实验数据，整体识别率达到了97.81%，表1中有两组实验室数据高斯核为 $K_1(13, 11, 25)$ 和 $K_2(25, 19, 23)$ 。高斯模糊的特性：模糊半径越大，图像越模糊。提取图像细节信息过程中，不同的模糊半径，影响高斯模糊的效果，提取的边缘细节信息不同。本文方法采用三尺度高斯模糊，不同模糊半径的模糊图像之间像素差值越大，提取的边缘信息越详细，人脸表情增强图像表达越明显。表1中两组参数的实验数据平均值相同，各个表情的识别率并不完全相同，两组参数对平均识别的影响相同，对每种表情的影响不同。因此，在高斯模糊不同的参数值下，提取的细节图像对表情的平均识别率可能产生相同影响，对每种表情的识别存在不同影响。

表1 高斯模糊半径取值的最优识别率(%)

高斯核	中性	愤怒	厌恶	恐惧	微笑	悲伤	惊讶	整体
K_1	98.00	99.33	100	96.67	100	94.67	96.00	97.81
K_2	96.00	98.67	100	95.33	100	96.67	98.00	97.81

高斯模糊的标准差也是影响效果的重要因素，为了探究了标准差对本文算法的影响，分别做了高斯核取值 K_1 和 K_2 下的实验。在标准差 σ 取区间[1, 25]的奇数值的实验中，高斯核取 K_1 时， σ_1 分别取(23, 21, 21)和(21, 21, 21)获得相同的测试数据，并且是该组测试下最优整体识别率。高斯核取 K_2 时， σ_2 取(25, 19, 23)获得该组测试下最优整体

识别率。如图9所示。高斯模糊是把正态分布应用于数字图像处理。高斯模糊的标准差是决定高斯模糊权值分布的重要变量。 σ_1 取两组不同的实验参数得到完全相同的实验数据，不同尺度的高斯模糊提取细节信息时，产生对图像的加强相同的强度，因此识别率相同；在3层高斯模糊提取细节过程中，在表1中的两组参数下，系统平均识别率得到提高，高斯模糊的标准差参数能够影响7种表情的识别率。 σ_1 和 σ_2 两组参数得到实验结果比较，厌恶和恐惧两种表情的识别率，前者略低于后者，但是，整体的平均识别率高于参数 σ_1 得到的实验结果。因此(25, 15), (19, 11), (23, 19)为多尺度高斯模糊的最优参数，使得系统识别率达到最高，系统平均识别率为98.19%。

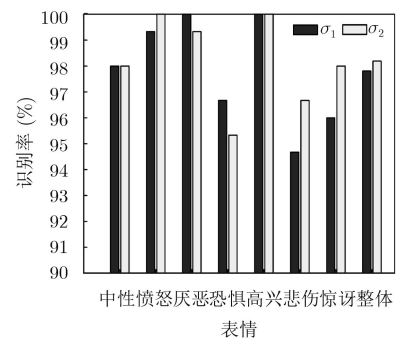


图9 实验结果比较

在CK+数据集中，7种表情的正确识别率和错误识别率分布见表2，7种表情的P-R曲线如图10所示。图10中，愤怒表情和厌恶表情的查准率为1，是一条直线；中性表情随着召回率的提升，查准率略低于其它表情。面部的情感表情主要通过面部器官的表达，当面部器官表达动作不明显，表情表达不明显，中性表情和其它表情的差异微小，容易造成错误识别，其中悲伤表情产生的面部部位动作不明显，容易产生混淆，系统将表情转换为利于系统计算的数据特征，数据特征差异不明显，系统对特征数据容易产生分类错误，因此，悲伤表情和惊讶

表2 表情识别率分布表(%)

	中性	愤怒	厌恶	恐惧	高兴	悲伤	惊讶
中性	98.00	0	0	0	0	2.00	0
愤怒	0	100.00	0	0	0	0	0
厌恶	0	0	99.33	0	0	0.67	0
恐惧	0	0	0	95.33	2.67	0	2.00
高兴	0	0	0	0	100	0	0
悲伤	3.33	0	0	0	0	96.67	0
惊讶	1.33	0	0	0.67	0	0	98.00

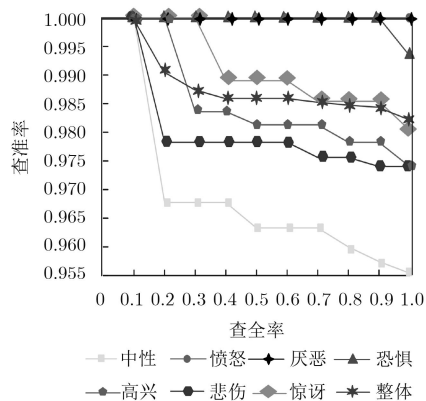


图 10 P-R曲线

表情的错误识别率集中体现在了中性表情, 中性表情错误识别率集中在悲伤表情; 尽管面部表达恐惧表情与高兴、惊讶等表情的面部动作不同, 产生的可视化的几何和纹理特征信息不同, 但是对于面部部位表达动作较大的这3种表情, 计算机系统将表情的表达特征转化为数据特征, 有些表情的数据特征具有相似性, 因此, 产生了错误识别。

3.3 实验结果比较

为了说明多尺度细节增强算法在表情识别上具有增强作用, 本文还做了另外两组实验, 表情图像识别方法是未采用图像金字塔和未采用细节增强的情感识别方法; 细节叠加识别方法是未采用图像金字塔且采用图像细节增强的情感识别。实验结果如图11所示。表情图像识别方法的表情识别率低于细节叠加识别方法的表情识别率, 平均表情识别率低于细节叠加识别方法4.28%, 细节叠加识别方法是在表情图像识别方法的基础上增加了图像细节增强计算, 提高了图像细节信息的表达能力, 人脸表情集中表现部位的特征更明显, 表情类别更容易区分, 识别率更高。多尺度细节叠加算法是在细节叠加识别方法基础上增加了图像金字塔。除恐惧表情和悲伤表情外, 多尺度细节增强算法的各个表情识别率均高于细节叠加识别方法识别率, 平均识别率

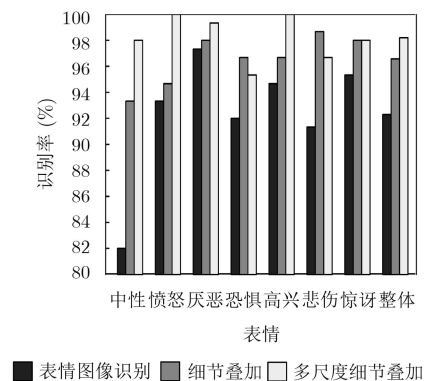


图 11 多尺度细节增强效果对比图

高出1.62%。图像金字塔对表情图像采样, 丰富了图像中表情的特征信息, 提高表情整体识别率。

表3是本文方法与CNN与LeNet-5^[15]方法比较。在小数据量样本中, 与CNN^[11]相比较, 其中表情的识别率中, 除了厌恶和惊讶两种表情略低于Lopes提出的CNN方法, 其它表情识别率和表情的整体识别率均具有更高的识别率, 本文方法与LeNet-5^[15]相比, 本文方法的表情识别率均较优。

表 3 本文方法与CNN与LeNet-5方法比较(%)

方法	中性	愤怒	厌恶	恐惧	微笑	悲伤	惊讶	整体
CNN ^[11]	95.15	91.11	99.44	92.00	100	82.14	98.80	95.75
LeNet-5 ^[15]	65.37	76.30	87.59	80.92	94.26	82.23	94.55	83.74
本文方法	98.00	100	99.33	95.33	100	96.67	98.00	98.19

不同方法整体识别率对比, 如表4所示。本文方法比Jia等人^[16]提出的基于LBP-Patches加权的多层稀疏表示的识别方法, 平均识别率高出了5.89%, 比Zhou等人^[17]提出的基于Gabor和NMF方法高出0.09%, 比基于卷积神经网络算法的文献^[11]也具有较优的结果。

表 4 不同方法识别率对比(%)

方法	识别率
LBP ^[16]	92.30
Gabor ^[17]	98.10
CNN ^[11]	95.75
F ² CNN ^[12]	96.20
LeNet-5 ^[15]	83.74
本文方法	98.19

本文从整体上和传统方法(LBP, Gabor)做了比较, 同时与卷积神经网络做了比较, 结论是本文方法相比深度学习方法, 省去了较为耗时的训练时间, 同时不过分依赖于样本数据, 同时获得了与基于深度学习方法略优的识别结果。

4 结束语

本文提出的多尺度细节增强人脸表情识别方法, 对图像取图像金字塔、图像细节增强对面部表情识别率的提高具有重要的作用。与其它方法相比, 本文的图像金字塔和细节增强, 能够在梯度特征计算过程中, 增强了不同表情特征描述子的差异, SVM分类器能够更好地训练分类模型, 提高了系统的识别率。

在人脸表情识别中, 人脸的点分布模型直接影响表情识别; 提取的特征描述子的数据维度过高, 会影响表情的识别效率。维度降低会降低表情识别

的正确率; 选取高斯模糊的参数做了大量的实验, 是非自动化实现。针对这些方面, 实现系统参数自动选择。在保证识别效果的前提下, 提高识别效率和获取人脸面部信息的完整性是下一步的研究工作。

参 考 文 献

- [1] MEHRABIAN A. Communication without words[J]. *Psychology Today*, 1968, 2(4): 53–56.
- [2] GUO Zhenhua, ZHANG Lei, and ZHANG D. A completed modeling of local binary pattern operator for texture classification[J]. *IEEE Transactions on Image Processing*, 2010, 19(6): 1657–1663. doi: [10.1109/TIP.2010.2044957](https://doi.org/10.1109/TIP.2010.2044957).
- [3] 王玮, 黄非非, 李见为, 等. 采用LBP金字塔的人脸描述与识别[J]. 计算机辅助设计与图形学学报, 2009, 21(1): 94–100, 106.
WANG Wei, HUANG Feifei, LI Jianwei, *et al.* Face description and recognition by LBP pyramid[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2009, 21(1): 94–100, 106.
- [4] 钟思志. 人脸面部表情识别算法研究[D]. [硕士论文], 华东师范大学, 2015.
ZHONG Sizhi. Research on facial expression recognition[D]. [Master dissertation], East China Normal University, 2015.
- [5] 杨凡, 张磊. 基于Gabor参数矩阵与改进Adaboost的人脸表情识别[J]. 计算机应用, 2014, 34(4): 1134–1138.
YANG Fan and ZHANG Lei. Facial expression recognition based on Gabor parameters matrix and improved Adaboost[J]. *Journal of Computer Applications*, 2014, 34(4): 1134–1138.
- [6] 童莹. 基于空间多尺度HOG特征的人脸表情识别方法[J]. 计算机工程与设计, 2014, 35(11): 3918–3922, 3979. doi: [10.3969/j.issn.1000-7024.2014.11.041](https://doi.org/10.3969/j.issn.1000-7024.2014.11.041).
TONG Ying. Facial expression recognition algorithm based on spatial multi-scaled HOG feature[J]. *Computer Engineering and Design*, 2014, 35(11): 3918–3922, 3979. doi: [10.3969/j.issn.1000-7024.2014.11.041](https://doi.org/10.3969/j.issn.1000-7024.2014.11.041).
- [7] TURAN C and LAM K M. Region-based feature fusion for facial-expression recognition[C]. Proceedings of 2014 IEEE International Conference on Image Processing, Paris, France, 2014: 5966–5970.
- [8] SAEED A, AL-HAMADI A, and NIESE R. The effectiveness of using geometrical features for facial expression recognition[C]. Proceedings of 2013 IEEE International Conference on Cybernetics, Lausanne, Switzerland, 2013: 122–127.
- [9] CHEN Junkai, CHEN Zenghai, CHI Zheru, *et al.* Facial expression recognition in video with multiple feature fusion[J]. *IEEE Transactions on Affective Computing*, 2018, 9(1): 38–50. doi: [10.1109/TAFFC.2016.2593719](https://doi.org/10.1109/TAFFC.2016.2593719).
- [10] 任福继, 于曼丽, 胡敏, 等. 融合表情和BVP生理信号的双模态视频情感识别[J]. 中国图象图形学报, 2018, 23(5): 688–697.
REN Fuji, YU Manli, HU Min, *et al.* Dual-modality video emotion recognition based on facial expression and BVP physiological signal[J]. *Journal of Image and Graphics*, 2018, 23(5): 688–697.
- [11] LOPES A T, DE AGUIAR E, DE SOUZA A F, *et al.* Facial expression recognition with convolutional neural networks: Coping with few data and the training sample order[J]. *Pattern Recognition*, 2017, 61: 610–628. doi: [10.1016/j.patcog.2016.07.026](https://doi.org/10.1016/j.patcog.2016.07.026).
- [12] ZHANG Chongsheng, WANG Pengyou, CHEN Ke, *et al.* Identity-aware convolutional neural networks for facial expression recognition[J]. *Journal of Systems Engineering and Electronics*, 2017, 28(4): 784–792.
- [13] LUCEY P, COHN J F, KANADE T, *et al.* The extended cohn-kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression[C]. Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, San Francisco, USA, 2010: 94–101.
- [14] YAO Yongqiang, HUANG Di, YANG Xudong, *et al.* Texture and geometry scattering representation-based facial expression recognition in 2D+3D videos[J]. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2018, 14(1S): 18.
- [15] 李勇, 林小竹, 蒋梦莹. 基于跨连接LeNet-5网络的面部表情识别[J]. 自动化学报, 2018, 44(1): 176–182.
LI Yong, LIN Xiaozhu, and JIANG Mengying. Facial expression recognition with cross-connect LeNet-5 network[J]. *Acta Automatica Sinica*, 2018, 44(1): 176–182.
- [16] JIA Qi, GAO Xinkai, GUO He, *et al.* Multi-layer sparse representation for weighted LBP-patches based facial expression recognition[J]. *Sensors*, 2015, 15(3): 6719–6739. doi: [10.3390/s150306719](https://doi.org/10.3390/s150306719).
- [17] ZHOU Jun, ZHANG Sue, MEI Hongyan, *et al.* A method of facial expression recognition based on Gabor and NMF[J]. *Pattern Recognition and Image Analysis*, 2016, 26(1): 119–124. doi: [10.1134/S1054661815040070](https://doi.org/10.1134/S1054661815040070).

谭小慧: 女, 1977年生, 博士, 副教授, 研究方向为计算机图形图像, 虚拟现实.

李昭伟: 男, 1992年生, 硕士生, 研究方向为情感计算.

樊亚春: 女, 1978年生, 博士, 副教授, 研究方向为计算机图形图像, 人机交互.